View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by UCL Discovery

Taylor & Francis Group

Comparative Education

ISSN: 0305-0068 (Print) 1360-0486 (Online) Journal homepage: http://www.tandfonline.com/loi/cced20

Negative capability? Measuring the unmeasurable in education

Elaine Unterhalter

To cite this article: Elaine Unterhalter (2017) Negative capability? Measuring the unmeasurable in education, Comparative Education, 53:1, 1-16, DOI: 10.1080/03050068.2017.1254945

To link to this article: http://dx.doi.org/10.1080/03050068.2017.1254945

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 24 Nov 2016.

Submit your article to this journal

Article views: 581

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=cced20

Download by: [UCL Library Services] **Date:** 21 February 2017, At: 08:22

COMPARATIVE EDUCATION, 2017
VOL. 53, NO. 1, 1–16
http://dx.doi.org/10.1080/03050068.2017.1254945

Routledge
Taylor & Francis Group

OPEN ACCESS

Negative capability? Measuring the unmeasurable in education

Elaine Unterhalter

University College London, Institute of Education, London, UK

ABSTRACT
This introductory article to the special issue of *Comparative Education* on measuring the unmeasurable in education considers measurement as reflecting facts and uncertainties. The notion of negative capability is used metaphorically to depict some limits of what is measurable, and portray aspects of the process of education, associated with uncertainty and public scrutiny of complexity. Four overarching questions – what, when, why and how – have guided the reflections of the authors who have contributed to the special issue. *What* are we measuring when we try to measure the unmeasurable in education and what are we not measuring? *When* have attempts been made to measure the unmeasurable in education, what metrics have been adopted in which contexts, and with what outcomes? *Why* have measures been adopted as indicators of the unmeasurable, such as human rights? *How* have particular historically located organisations approached the problem of measuring the apparently unmeasurable in education, with what epistemological, normative and conceptual resources, and consequences? The introductory article looks at measurement as a form of negative capability in some discussions of history of social statistics in education, the current debate over indicators for the Sustainable Development Goals, and how to measure gender equality in education.

KEYWORDS
Education measurement; capability approach; SDGs; gender equality

In 1817, the English poet, John Keats, wrote a letter to his brother in which he mused on the nature of Shakespeare's talent. He used the phrase 'negative capability' to describe 'when a man is capable of being in uncertainties, mysteries, doubts, without any irritable reaching after fact and reason'. Keats described a process of edging towards something creative, but not easily definable. The concept of negative capability has had considerable influence in work on psychoanalysis and literary criticism. It has potential to be generative of ways to think about education, as a process of being and becoming. The notion of negative capability captures features of education practice that reflect both facts and uncertainties. But interpreting education in this way goes against a dominant trend in research, which leans towards clarity on what works, and precision with regard to how this should be measured. This special issue explores how we can understand measurement in areas of education policy, planning and practice that have not previously been

CONTACT Elaine Unterhalter e.unterhalter@ucl.ac.uk

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

considered measurable. It thus considers the potential of measurement itself as a form of negative capability.

The discussion in this special issue steps off from the premise that education is not just one thing, for example, a learning outcome, linked to performance in a test, or the numbers enrolled in a particular school phase. Many aspects of education defy measurement. Educational relationships that are social, emotional, epistemological, normative, political, cultural and economic cannot be simply measured. Thus many writers on education are critical of the political ways in which measurement as a method in comparative analysis has been used (Cowen 2014; Gorur 2014; Morris 2015; Auld and Morris 2016). The rigid application of some methods, such as randomised control tests, reduces complex information to apparently simplified causal relationships. Numerical data on education have come to play a key role in describing education systems, and prescribing reform inspired by new public management, as a number of historical studies show (Lawn 2013; Meyer and Benavot 2013; Goldstein and Moss 2014). However, neither the historical work, which raises questions about the interpretation of data, nor the questions posed by the critical policy literature has dented a concern to define education change as science linked with a selection of facts. Unmeasurable processes in education are routinely addressed through appeals to measurement or indicators. The Millennium Development Goals (MDGs) of 2000–2015, the Sustainable Development Goals (SDGs), adopted in 2015, PISA and university rankings are some of the most well-known projects that attempt to measure aspects of education, linking the precision of measurement, with imprecisely formulated values.

In these instances, numbers, and numeric forms of interaction, are often used to stand as proxies for complex relationships that are really unmeasurable. Some highly salient features of education, which change over time, place and from different personal perspectives, entail understandings of wellbeing, agency, tacit and applied knowledge, criticality, creativity, equality and public good. All of these entail social relationships which interact with contexts. These multiplicities are unmeasurable in numeric form. The precision claimed for measurement may actually obscure the importance of what is not measured. There is thus a tension between what is easily measurable, but may not be significant, and what is of major importance, but cannot be measured. However, merely asserting this binary suggests it is impossible to translate between the different frameworks. In earlier work, I have argued for developing a transversal dialogue between approaches to education that map what works, and those that argue for what matters (Unterhalter 2009). In this introductory article to the special issue I consider measurement of a range of unmeasurable issues as a form of creative negative capability. This acknowledges some of the limits of what is measurable, and keeps open the dynamics of uncertainty and public scrutiny of complexity.

A couple of issues frame my concern with this negative capability. Firstly, what may be possible and desirable in measuring the unmeasurable in education at particular moments? Granted that some attenuation of the notion of comparison will always be associated with exercises to develop metrics or indicators, is it feasible or desirable, to engage critically with this process? Is the use of education data always linked with versions of new public management and deformations of wider and deeper values articulated through education? Or can a more reflective approach to measurement be associated with a more democratic process concerning education data? Secondly, what are the

contextual politics of projects concerned with measuring the unmeasurable in education? What happens to normative concerns of education with equality, wellbeing, agency and public good when projects are undertaken to measure these, which acknowledge some of the limits of measurement? Are meanings for these values expanded or diminished by the use of metrics? Whose interests are advanced, and whose neglected? May failures of precision in measurement also be associated with gains in knowledge, advocacy and more critical and reflective practice?

Four overarching questions – what, when, why and how – have guided the reflections of the authors who have contributed to the special issue. I consider these in my discussion in this introductory article. *What are* we measuring when we try to measure the unmeasurable in education and what are we not measuring? *When* have attempts been made to measure the unmeasurable in education, what metrics have been adopted in which contexts, and with what outcomes? Thus when states and organisations have used numbers as proxies for values or relationships, is it always associated with forms of domination and regulation, or have there been contradictory dynamics? Has a demand for measurement ever been a demand of subaltern voices? *Why* have measures been adopted as indicators of the unmeasurable, such as human rights? What debates accompanied these discussions, and how did technologies associated with data link with some of the decisions taken? *How* have particular historically located organisations approached the problem of measuring the apparently unmeasurable in education, with what epistemological, normative and conceptual resources, and with what consequences?

The articles which follow all address aspects of these questions. In this introductory article, I explore the notion of measuring the unmeasurable in education as a form of negative capability looking at some instances which exemplify these questions. In the analysis below, I look at some discussions of history of social statistics in education, considering what aspects of the unmeasurable are measured and not measured, in which contexts and with what outcomes. I go on to look at the current debate over indicators for the SDGs exploring how particular organisations have engaged with the potential and limitations of measuring the unmeasurable, and what possibilities for further work in this area may exist. Discussions of how to measure gender equality in education are examined as an instance of why indicators have been considered by advocates of equalities and some of the debates that have accompanied this.

Analysing measurement and statistics in education

Negative capability which connects what is measured with what is valued and not easily quantified is a feature of the history of the use of statistics in education. A number of historical accounts of what has been measured in education show how the arguments made about measurement were linked with ideas about improving education quality, policy delivery and enhancing accountability (Lawn 2013; Meyer and Benavot 2013; Goldstein and Moss 2014). The turn to data in education can be situated in a wider literature which maps enhancing evidence to underpin the making of policy, attempting to ensure processes that are more democratic and transparent (Davis, Kingsbury, and Merry 2012).

But difficulties abound in engaging with measurement as a form of negative capability. It is rare for organisations or analysts who develop indicators to keep both a critical

perspective on what is measured, and engage in substantive participatory discussions about this with the people who are most affected and to whom they are accountable. As contributors (Fukuda-Parr, Yamin, and Greenstein 2014; Rottenburg et al. 2015) to a number of critical discussions show, the processes associated with the selection of indicators and the manipulation of data are generally not open or subject to critical review. Education is no exception. Indicators, these commentators point out, have become part of a discourse of regulation linked to new public management, rather than a process to enhance democratic participation and review of decision-making. Some writers suggest that the development of any indicator disorientates the complex process of analysis for social change. A number of writers on women's rights and gender equality have been critical of indicator projects for distracting from the persistent injustices and violence associated with deeply entrenched inequalities and the wide range of social and political activism needed to transform this (Antrobus 2006; Baksh and Harcourt 2015). A further line of critique, powerfully voiced through theorists of postcoloniality (e.g. de Sousa Santos 2015) is that indicators as a form of governance and accountability fail to confront the real and pressing contemporary issues around global inequalities, dispossession, war, racism and violence. Alexander (2015), commenting on proposed indicators for the SDGs, critiqued the conflation of measurement and indicators pointing out some perverse effects on how pedagogy is understood and evaluated. Education indicator discussions, he concluded, had been particularly narrow-minded and neglectful of contexts.

Thus negative capability associated with measuring the unmeasurable may be simply negative, and not linked in any way with capability. Some of the analysis of the use of measurement in education is linked with the emergence of technologies to count, store and analyse data (Crossley 2014). A number of critiques question how this is done and highlight some of the consequences of imperfect processes. Data collection linked with state formation, when there are not democratic processes to build institutions and critique data, often generate perverse consequences. Thus states, which represent elites, or particular racial or ethnic groups, generally do not consider how relationships of power affect the collection of census data (Kertzer and Arel 2002; Simon and Piché 2012). A number of studies of the anthropology of how data is collected for surveys, draw out problems around defining and categorising households and the assumptions about relationships that are made (Randall and Coast 2015; Oya 2015). Carr-Hill (2013) notes how the identity of census enumerators has a bearing on underestimations of children out of school and poverty, as enumerators do not go to some of the most difficult areas.

Whether data and education statistics on the unmeasurable are negative or a site of negative capability is illuminated in the experience with the MDGs. These used measurement of targets and indicators as a form of development regulation, but these measures were also linked to some values about equalities, rights and agency, all of which were unmeasurable in this way. Thus, for example, MDG 3 was concerned with the wide terrain of gender equality and the empowerment of women. The associated targets and indicators were much narrower, using a measure of gender parity, that is, the ratio of women to men enrolled at different levels of education. There was no assessment of women's engagement in other fields of activity (Waage et al. 2010; Sen and Mukherjee 2014). This misalignment generated a literature which brought together reflections on the history of social statistics and varied experiences with measuring the unmeasurable

across social development sectors. In some cases, the misalignment had perverse effects, while in others it mobilised advocacy campaigns (Unterhalter 2014a; Fredman, Kuosmanen, and Campbell 2016). The *Power of Numbers* research project, co-directed by Sakiko Fukuda Parr and Alicia Yamin, illuminated how the history of using numbers to measure development was different for different social sectors involved with the MDGs (Fukuda-Parr, Yamin, and Greenstein 2014). Many MDGs, including education, had a much wider hinterland of debates around measurement than the MDG framework gave credit for (Unterhalter 2014a; Barrett 2016). The authority in the MDGs sometimes eclipsed these other processes of debate and critique.

But this terrain of critical discussion can sometimes be extremely generative. The negative capability of measuring the unmeasurable has sometimes been linked with a more nuanced approach to method, which seeks to go beyond the divides between quantitative and qualitative approaches. Qualitative Comparative Analysis is emerging as a methodology which complements established quantitative and qualitative methods in education, because it seeks to look at relationships of causation, rather than more abstract kinds of connection (Marx, Rihoux, and Ragin 2014). Realist evaluation is a theory-based evaluation approach that investigates how and why particular programmes work or do not work, for whom, and in what circumstances (Pawson and Tilley 1997). The approach is currently being used in the Building Capacity to use Research Evidence programme, which is yielding some promising fine-grained results (Punton, Vogel, and Lloyd 2016). Work using methods such as these could inform thinking about negative capability and measurement.

Measuring the unmeasurable has been linked to a range of work around equalities and rights. The intellectual community that coheres around the Human Development and Capability Association (HDCA) is one highly visible network working in this way. It draws on the analytical framing provided by the capability approach, as formulated in writings by Sen (1999) and Nussbaum (2011) and has generated significant critical engagement with the idea that measuring capabilities, entails not so much counting outputs, such as numbers of children in school, but opportunities, that is capabilities, the conditions that constrain or facilitate these, and whether or not capabilities are realised as functionings. These theoretical resources have been associated with a number of innovative approaches to measuring the unmeasurable. Sen formulated the Human Development Index (HDI) in the early 1990s, and this work has been refined and expanded in 25 years of reports by the United Nations Development Programme (UNDP 2015). Other adaptations of the capability approach used in measuring social development are evident in the Multi-dimensional poverty index, formulated by Alkire, Foster and colleagues at the Oxford Poverty and Human Development Initiative (Alkire et al. 2015), the social institutions and gender (SIGI) index looking at institutions associated with underpinning gender inequality and equality, developed by Stephan Klasen and colleagues at OECD (Branisa et al. 2014), the indicators used for measuring equalities in the UK formulated by Burchardt and Vizard (2011) and the SERF index looking at socio-economic inequalities between countries and regions (Fukuda-Parr, Lawson-Remer, and Randolph 2015). These are all highly creative initiatives to understand the space of capabilities, understood as either institutional, or the outcome of overlapping social relations across sectors. However, all draw on existing datasets of education, that is numbers enrolled, or levels of attainment. The complexities of relationships within or through education are not yet fully captured by these

approaches to measure at national and sub-national levels. They generally do not acknowledge the limitations of the education data they draw on. In addition, with the exception of the approach developed by Burchardt and Vizard (2011), there has not been much participatory engagement in reviewing these approaches to measuring equalities or addressing poverty.

The literature on education measurement suggests both possibilities and difficulties associated with measuring the unmeasurable. Evaluations of this negative capability entail posing questions regarding what is measured and how. This discussion has to be put in the context of hierarchies in access to and use of data, and participation in assessing metrics in particular countries. Whether indicators, evaluations or measurement enhance forms of accountability and public debate, or are irrelevant to particular contestations, are empirical questions still awaiting documentation.

Measuring the unmeasurable in the SDGs

Depending on one's perspective, the SDGs represent either a watering down of an international approach to addressing substantive international human rights commitments, built up since the 1970s, or an opportunity to take these further (Gore 2015; Scheyvens, Banks, and Hughes 2016; Pogge and Sengupta 2016). Education has prominence in the SDGs, both as a substantial goal with a number of targets, and as a component of other goals. Thus this policy framework presents considerable opportunities as well as risks. We can view the SDGs as a site of negative capability because they suggest an enhanced access to resources, networks, ideas and centres of debate, not exclusively tied to the controversial measurement agenda. But the challenge of realising this negative capability partly rests on understanding how particularly located organisations utilise this space.

Within education circles, the different perspectives on the SDGs have some distinctive connotations which distinguish this field of discussion from other policy areas. Some social activists, who work, for example, on women's empowerment or urban environments, are highly critical of the SDG agenda, which has excluded many long held aspirations (Barnett and Parnell 2016; Esquivel 2016). However, in education, the SDG agenda not only expanded the focus of the education MDGs, it also goes considerably beyond the vision laid out in 2000 for Education for All (EFA). The MDGs had focussed on access and completion of primary education, and gender parity at all levels of schooling. They had been silent on what was learned in school, and ignored all other phases and sites of education (Mundy and Manion 2015; Barrett 2016). While the 2000 Dakar Platform on EFA had a wider range of goals than the MDGs, these too were vague around education quality, and said nothing on post compulsory phases. The SDGs, by contrast have targets that relate to all phases of education (UN 2015).

In contrast to the history of MDGs, which were formulated by a small group of experts sitting in New York (Vandemoortele 2011), the SDGs have been associated with a greater level of participation in suggesting and commenting on goals, targets and indicators (Sénit, Biermann, and Kalfagianni 2016; Sayed 2013). In the period (2014–2016) leading up to the selection of targets and indicators, and in the months that followed the adoption of the SDG policy text at the UN, advocacy groups used an emerging concern with measurement to argue for education indicators that could convey situations around disability, gender, human rights and early childhood development, which had previously

not been included in the MDGs or EFA (Tabbush 2014; Unterhalter 2015; Antoninis, Delprato, and Benavot 2016; Brolan 2016). This is an instance of the SDGs as a site of negative capability generating demands around substantive education equality issues, using the language of measurement,

SDG 4 acknowledges the importance of attending to quality education at all levels from pre-primary to technical and higher education. The goal includes targets on developing understandings of equalities, citizenship and sustainability (Target 4.7). Other SDGs clearly require education processes, as the most recent issue of UNESCO's *Global Education Monitor* (UNESCO 2016) sets out. For example, SDG 1 is concerned with ending poverty and creating the social policy environment in which poverty can be addressed. It thus invites attention to the knowledgeable practice of teachers as a key resource for helping to support the poorest children on pathways beyond poverty and inequality, ensuring education is delivered in ways that are gender equitable, and supportive to children, irrespective of background. Thinking about developing appropriate indicators for teacher practice in this area, using participatory processes, that take account of context and individual uncertainties, may enhance creative ways of engaging with this process. SDG 5 is concerned with ending discrimination against girls and women, eliminating violence, including school-related, gender-based violence (SRGBV), and supporting empowerment. The targets for this goal are concerned with networks of practice to eliminate violence. A recent rigorous literature review of interventions on SRGBV, notes major evidence gaps exist on how to provide safe, inclusive and violence-free learning environments for girls and boys. The research that has been done has tended to be skewed towards evaluations of short-term interventions at a moment of practice, with little long-term follow-up. (Parkes et al. 2016). This may be an instance of what is unmeasurable, but it also suggests that some critical examination of whether to develop indicators in this field may help build discussion and investigation in this area.

The SDGs are clearly a site of political and policy contestation. They have generated a context in which debates can be conducted as to why the unmeasurable should or should not be measured, and what some of the consequences of measuring in this way might be. Although they are often discussed as examples of regulation and control, they also exemplify negative capability.

The debate about measuring gender in education

Measuring gender equality in education is a feature of many of the most well-known metrics and the question of what is measured, why and how surfaces very clearly in this area. What is easy to measure, numbers of girls and boys, does not always express what is important to measure. The concept of gender equality in education may be unmeasurable, but debates as to how to measure this have been very illuminating, indicating some generative dimensions of negative capability.

For many decades the standard measure of gender equality in education has been gender parity. This entails expressing the number of girls enrolling, progressing or achieving in schools as a proportion of the number of boys. Gender parity is used in PISA data, indicators for the education and gender MDGs, and some of the SDGs. Gender parity also appears in the UNDP Gender Development Index. However, this widespread use of this measure, ignores extensive scholarly and practitioner discussion that the meanings of

gender, equality, and education are not as settled and simple as the use of the standard measure of gender parity would suggest (Sen and Mukherjee 2014; Unterhalter 2014b; DeJaeghere 2015). Gender equality in education is linked with wellbeing, agency, aspects of embodiment and lack of violence, knowledge and criticality, public good, social relationships and context. Thus the notion is a rich and complex one, which suggests substantive relationships, not reducible to or expressed through gender parity.

Gender parity as the measure of gender equality in schooling was widely promoted in key World Bank publications of the early 1990s (King and Hill 1993) and used in the UNESCO Global Monitoring Reports produced from 2003 to review progress on EFA. However, gender parity does not sufficiently capture the range of relationships and values associated with the notion of gender equality in education, and what learning outcomes relating to gender equality might entail. The technique of measuring gender parity tends to underplay a connection between education, women's rights and social justice, and thus provides inadequate information to evaluate progress against the gender equality aspirations of Constitutions or international policy texts, such as Education 2030. Measuring gender parity alone thus does not help policy and strategy build towards the development of substantive equality because it is not a clear enough indicator of the relationships within and beyond education that need to be changed to achieve this.

A number of difficulties about measuring education, equalities, justice and wellbeing collide in the problems of measuring gender equality and inequality in education which echoes the questions posed to the authors contributing to this special issue.

Firstly are we measuring *what* certain groups defined by binary gender categories of women and men do or do not have? In other words, are we measuring resources noting whether they are male or female (numbers of teachers, images in textbooks), amounts spent on girls or boys, or objective states of levels of enrolment, attendance or attainment in education for different groups of children? What we measure codifies a certain meaning of gender, while other aspects associated with the concept may be unmeasurable, for example, features of gender relations, sexualities and aspects of power.

Secondly, *why* are we interested in measuring gender equality in education? Are these concerns driven narrowly by notions of efficiency and good planning linked with governments or other kinds of large-scale organisation? Or are concerns with measuring gender equality in education driven by wider normative concerns and organisations that give considerable space to consultation, participation and attending to those who have been excluded. If organisations are concerned with expanding the capabilities of what people can do or be, do processes of consultation and reflection bridge the assessments of researcher, technicians, practitioners and other participants in developing appropriate indicators? Do we want these metrics to provide some indicator of what is right or good, and what normative language helps establish this?

Thirdly, if we acknowledge that education and gender equality in some way signal a relationship between subjective and objective conditions, given that the notion of gender knits together descriptive, interpretative and normative concerns, what kind of metric allows us to consider this? How do we measure, and with what methods? What are the processes that can be angled to understand the relationships and uncertainties between contexts, emotions, gendered relations, and outcomes?

The history of measuring amounts of education resources – for example, numbers of classrooms, teachers, pupil/teacher ratios or numbers of pupils progressing at each

level – is useful in planning and managing distribution, but becomes very difficult when trying to measure what it is that people value about education, and how this connects with other dimensions of wellbeing, such as critical questioning, gender equality or public good. This is not to say studies of these mental states have not been made and some international assessments, for example, for PISA, measure aspects of autonomy and civic mindedness, but, the implications of these studies are more diffuse than those that are more limited and focus only on the distribution of amounts of education particular groups have. The OECD study of gender and education, with its stress on ABC (Aptitude, Behaviour, Confidence) attempts to address an aspect of the link between subjective and objective states and an aspect of what we value about education beyond time in school (OECD 2015). However, this has not been discussed as widely as PISA and does not have impact of the PISA data on national education systems.

If we are concerned with measurements of 'what', such as resources for education, or amounts of education received, gender equality comes to be defined as groups of girls and boys, women and men, who have equivalent amounts of a particular resource. It is for this reason that gender parity remains such a key thread in international policy discussions. But measuring the distribution of resources using gender parity as the metric has long been criticised. Critiques range from the importance of attending to the texture of gendered experiences and the multiple registers of how people talk (Henderson 2015). Studies draw out the diversity of moments associated with gendered experiences of education, for example, the complexity of pedagogic encounters, or forms re-contextualisation (Chisamya et al. 2012). The importance of understanding vulnerabilities, forms of violence, opportunities and constraints (Parkes 2015) far exceed what gender parity can tell us. In addition, the simple binary of opposing men and women does not allow us to account for categories of gendered identity (trans- or intersex) that cross or confound these divisions. Gender parity is critiqued for being heteronormative (Mishra 2016). These debates about the limitations of gender parity, bring out how the stress on amounts and issues of planning, fails to express the social justice aspirations of those concerned with the multi-dimensionality of the ideas of gender equality in education. This echoes some of the wider literature on how we measure justice, equality, empowerment and wellbeing (Kabeer 1999; Brighouse and Robeyns 2010; Clark 2014). Thus it is illuminating to draw on some of to try to unravel some of the problems of measuring gender equality in education.

Clark (2014) identifies three major approaches to measuring wellbeing, distinguishing those that focus on utility, and actual or informed desires, those that focus on resources, and those that rest on list-based approaches. Implicit in the work of the critics of the gender parity measure is the notion that gender equality in education goes beyond equal amounts. Thus in some way, it is a part of wellbeing, happiness or some other aspect of quality, equality or justice. Thus its 'goodness' extends beyond parity of amounts between men and women and its 'rightness' or sense of justice is not simply met by noting this basic pattern of distribution. However, it is one step for critics of gender parity or the simple measure of the distribution to say that it is not an adequate measure of gender equality in education. It is an altogether more difficult proposition for this group of critics to say what measuring gender equality in education should be.

Some important considerations are that measuring gender equality in education should have something to do with taking seriously aspects of agency, particularly, given

the history of women's exclusion, subordination and injustice, noted in every country in the world, the autonomy and voice of women. This includes the right to articulate, review and have perspectives on justifications for social arrangements and the complexities and unevenness of this process. It should also include some assessment of the conditions, associated with political economy and socio-cultural practices that maintain injustices against groups, or classes marked by particular gendered dynamics. The metric should also in some way allow for some evaluation of the strength of processes for change in the direction of greater substantive gender equality and justice. Lastly, because of the significance of bodily and emotional vulnerability in many meanings of gender, measuring or evaluating gender equality in education should signal some protection of bodily integrity and concern with emotional support. Some of this work has been done in small scale studies looking at measuring aspects of gender empowerment locally, but there are difficulties in taking this to scale.

A number of what Clark (2014) has called list-based approaches to measurement of rights, capabilities or human development attempt to overcome the difficulties utility and resource-based approaches to measuring gender equality and education encounter. But the problem with list-based approaches, as Sen (2005), generator of one of the most famous of such approaches, the HDI, acknowledges is that the areas on the list and the weights given are always arbitrary. Thus list-based approaches, in trying to correct for the stress on revealed desires, which dogs utilitarianism, come to suggest a particular formulation of what human development, or gender equality or education quality is. List-based approaches will always be open to the critique that the people who developed the list and gathered the information were not those who expressed views regarding what they could or could not do and be with regard to education or gender equality. This criticism is often levelled (e.g. Sen 2004; Robeyns 2005; Walby 2011) at Nussbaum's (2000) list of central human capabilities which goes way beyond the HDI in including aspects of bodily integrity and emotions, more familiar to discussions of gender equality, than those that focus only on economics. Robeyns (2003) has proposed a method for developing lists that is open to public scrutiny.

There have been a number of list-based approaches developed in education, as Clark discusses in his article in this issue. None of these list-based approaches have addressed the question of measurement for national or international policy, and only Loots and Walker (2015) have developed a list which considers gender issues, but only in higher education. In the wider gender equality literature there are a number of list-based approaches (UNDP Gender Inequality index, OHPI Women's Empowerment in Education index, World Economic Forum Gender Gap, SIGI index), but these either do not address education, or take only aspects of the administrative data on gender parity in education as their education component. This approach is taken by the Gender Gap index, but if gender parity is the point of contention, repositioning this on the list does not get over the problem of the narrow range of resource-based issues it measures. List-based approaches regarding aspects of human wellbeing have tried to signal some connection between subjective and objective conditions, and the list-based approaches in education, even if they do not address gender explicitly, have brought to the fore the question of the importance of subjective articulations. List-based approaches which do take gender seriously have gravitated either to the pole of revealed preferences, that is what learners or teachers say – for example, the work on the Care index (Miske, Meagher, and

DeJaeghere 2010) – or the pole of institutional claims, that is what laws or policy say, for example, the OECD SIGI index (Branisa, Klasen, and Ziegler 2013). Some work on equalities in the UK has tried to navigate between the two (Burchardt and Vizard 2007) but this work has dealt more with the provision of health and housing, rather than education, and engagements with gender are quite implicit. Dejaeghere (2015) notes several shortcomings of list based approaches. They do not express how capabilities are dynamically related to social arrangements. They are quite static and cannot fully express agency, aspiration or reflection on context. The lists might thus express aspects of capabilities, but not some of the creative tension of negative capability or all the dimensions of the capability approach, as mapped by Robeyns (2016).

The issue of what a measure of gender equality in education that goes beyond gender parity could look like, remains a fertile field of engagement. A number of SDG4 targets regarding access, progression and learning outcomes will be measured using an indicator of gender parity (UNECOSOC 2016). But the measure of gender equality in education associated with SDG4 are not yet completely settled. Thus there remains an opportunity to critically examine how we measure and track gender equality in education, its enabling environment and associated learning processes. In 2015, I suggested an indicator framework that might be able to do this (Unterhalter 2015) and some discussion with key individuals in the indicator policy community have followed.

However, to understand the relationships of gender equality, education and learning outcomes needs some careful conceptual work, which takes on board debates around pedagogies, policies and social context. In my work on a possible indicator I suggested the learning outcomes, however defined, should not be separated from the pedagogies that realise this and that the indicator should attempt to distil this. Some key domains I suggest to evaluate in relation to gender equality include discussion of gender and equality issues in curriculum and learning materials, opportunities for discussion of gender in teaching and learning, and strategies for correcting gender bias in pedagogies. We thus need to disentangle learning about gender, learning how to do gender equality and learning how to be gender equitable, and relate these to strategies and work with practitioners to build an enhanced indicator frame.

This review of some of the history of measuring gender equality in education reveals both the closures and openings evident in engagements with the negative capability of measuring the unmeasurable. The spaces opened up through the conceptual language of the capability approach allowed for a critical engagement with measuring the unmeasurable. Regardless of whether an enhanced indicator frame on gender is or is not adopted as part of the SDG process, the scholarly engagement with this question has helped to build connections and conceptual language that would not have come into being without this debate. A dialectical opposition between measuring and not measuring might have prevented this new exploration emerging. However, the negative capability associated with measuring the unmeasurable has been fruitful.

Engagements with measuring the unmeasurable

The articles which follow all engage with one or more of the questions I have posed about measuring the unmeasurable in education. Heinz Dieter Meyer shows that the search for precision in education measurement, and disputes about whether this captures what we

value in education, does not start with new public management or PISA. The debate has roots that reach across centuries. The Aristotlean notion of *phronesis* is useful in clarifying some of the unmeasurable facets of educational practice. Damiano Matasci traces the history of UNESCO's use of statistics to measure literacy in the 1940s and 1950s. He shows how rights were central to the organisation's orientation. At a time when colonialism was still a major force and contested some of the focus on rights and equalities, UNESCO sought to gain legitimacy through its use of what were claimed to be scientific measures. David Clark draws on a review of the literature on the capability approach and education. He evaluates a number of education capability lists, and contrasts what they detail with some data gathered from surveys he conducted in poor communities in South Africa. He suggests it may be possible to embrace the complexity and imprecision involved in measuring education and responding to the aspirations of the poor, sketcheing a methodological framework for this. Emily Hannum, Ran Liu and Andrea Alvarado Urbina discuss the ways in which poverty has been considered measurable in a number of cross national education surveys, such as PISA, Southern and Eastern African Consortium for Monitoring Education Quality (SACMEQ) and the Trends in International Mathematics and Science Study (TIMSS). They draw out some of the problems with the questions asked and the responses received, and suggest some methodological steps that are attentive to the vulnerabilities of poor children, and their ways of knowing about education and deprivation. Niall Winters, Martin Oliver and Laurenz Lange show how efforts to use new technologies to enhance training delivered to health workers have tended to be linked with metrics of accountability and scale. These do not take seriously the challenge of measuring better, which requires an appreciation of the socio-cultural dynamics of learning or the importance of connection to the health system. Nelly Stromquist documents new measures of academic outputs introduced in US universities in the last decade. She shows how the metrics have divided an academic community, separating research producers from teachers and leaving the question of how to evaluate teaching adrift from how research is measured. In the final article in this collection, Stephanie Allais looks at indicators associated with measuring how higher education contributes to society. She shows how the precision associated with current forms of measurement does not address key questions needed in planning what proportion of public resource should be spent on universities in countries such as South Africa, addressing multiple inequalities. She argues, nonetheless, that acknowledging what is currently unmeasurable but significant about the social value of higher education can help take discussion into productive terrain. Overall all the articles in this special issue show that the questions of measuring the unmeasurable is being actively and critically canvassed linked to particular terrains associated with inequality, injustice, and the contemporary complexities and uncertainties of work in education.

This introductory article has considered the process of measuring the unmeasurable in education as a form of creative negative capability. The discussion has shown that an acknowledgement of the limits of what is measurable has been an element of the history of debates around the use of education statistics, the SDGs and measuring gender equality in education. But in each area, I have shown it has been possible, under certain conditions, to keep open the dynamics of uncertainty and public scrutiny of complexity. Whether it is possible and desirable to measure the unmeasurable in education is not pre-given. There are examples that show that at particular moments this is a

site of negative capability, but there are other examples that show perverse effects or forms of closure. However, I consider a reflective approach to measuring the unmeasurable in education does enhance democratic, critical discussions about education data. Opening up normative concerns of education with equality, wellbeing, agency and public good to the possibilities of measuring the unmeasurable, have generally helped expand understandings of these concepts. This may not always be the case when these debates come to be co-opted by powerful organisations with ambiguous agendas. Thus concerns will always remain to consider whose interests are advanced, and whose neglected in discussions of measuring the unmeasurable. I have thus argued for work on measuring the unmeasurable as a form of negative capability that can support gains in knowledge, advocacy and more critical and reflective practice. It is a richly documented analysis, like the articles that follow, that will help build our insight as to whether this is indeed a productive form of uncertainty.

Acknowledgements

The ideas explored in this article have been discussed over a number of years with colleagues collaborating on panels presented at CIES and HDCA conferences, the CECEPS seminar series on education data, which I co-directed with Gemma Moss in 2011–2012, and the Power of Numbers research project, led by Sakiko Fukuda Parr and Alicia Yamin (2013–2014). I am very grateful to all who have given feedback in these fora on my evolving ideas. Warmest thanks to Bob Cowen and Joan Dejaeghere for very insightful comments on drafts of this paper.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

The workshop co-organised by UNGEI, FHI360, and the Institute of Education, in London, September 2015, co-ordinated by Nora Fyles, Shirley Miske and I, with support from the Macarthur Foundation, was an opportunity to look at measuring gender equality in education.

Notes on contributor

Elaine Unterhalter is Professor of Education and International Development, University College London, Institute of Education. She has worked on research projects on education equalities in a number of countries in Africa and published widely on comparative and international education. Her books include *Gender, Schooling and Global Social Justice* (Routledge), a co-edited volume on the capability approach and education (with Melanie Walker), several volumes on gender and education (with Sheila Aikman), and an introductory volume on education and international development (with Tristan McCowan). She is on the Editorial Board of *Comparative Education*.

References

Alexander, R. 2015. "Teaching and Learning for all? The Quality Imperative Revisited." *International Journal of Educational Development* 40: 250–258.

Alkire, S., J. Foster, S. Seth, M. Santos, J. Roche, and P. Ballon. 2015. *Multidimensional Poverty Measurement and Analysis*. Oxford: Oxford University Press.

Antrobus, P. 2006. "Gender Equality in the New Millennium: Goal or Gimmick?" *Caribbean Quarterly* 52 (2/3): 39–50.

Antoninis, M., M. Delprato, and A. Benavot. 2016. "Inequality in Education: The Challenge of Measurement." World Social Science Report, 63–67. Paris: UNESCO. http://unesdoc.unesco.org/images/0024/002459/245937e.pdf

Auld, E., and P. Morris. 2016. "PISA, Policy and Persuasion: Translating Complex Conditions Into Education 'Best Practice'." *Comparative Education* 52 (2): 202–229.

Baksh, R., and W. Harcourt, eds. 2015. *The Oxford Handbook of Transnational Feminist Movements*. Oxford: Oxford University Press.

Barnett, C., and S. Parnell. 2016. "Ideas, Implementation and Indicators: Epistemologies of the Post-2015 Urban Agenda." *Environment and Urbanization* 28 (1): 87–98.

Barrett, A. 2016. "Measuring Learning Outcomes and Education for Sustainable Development." In *The Global Testing Culture*, edited by W. Smith, 101–114. Oxford: Symposium Books.

Branisa, B., S. Klasen, and M. Ziegler. 2013. Gender Inequality in Social Institutions and Gendered Development Outcomes. *World Development* 45: 252–268.

Branisa, B., S. Klasen, M. Ziegler, D. Drechsler, and J. Jütting. 2014. "The Institutional Basis of Gender Inequality: The Social Institutions and Gender Index (SIGI)." *Feminist Economics* 20 (2): 29–64.

Burchardt, T., and P. Vizard. 2007. "Developing a Capability List: Final Recommendations of the Equalities Review Steering Group on Measurement." CASE Research Paper No, 121. London: London School of Economics.

Brighouse, H., and I. Robeyns, eds. 2010. *Measuring Justice: Primary Goods and Capabilities*. Cambridge, MA: Cambridge University Press.

Brolan, C. 2016. "A Word of Caution: Human Rights, Disability, and Implementation of the Post-2015 Sustainable Development Goals." *Laws* 5 (2): 22.

Burchardt, T., and P. Vizard. 2011. "Operationalizing'the Capability Approach as A Basis for Equality and Human Rights Monitoring in Twenty-First-Century Britain." *Journal of Human Development and Capabilities* 12 (1): 91–119.

Carr-Hill, R. 2013. "Missing Millions and Measuring Development Progress." *World Development* 46: 30–44.

Chisamya, G., J. DeJaeghere, N. Kendall, and M. Khan. 2012. "Gender and Education for All: Progress and Problems in Achieving Gender Equity." *International Journal of Educational Development* 32 (6): 743–755.

Clark, D. 2014. "Defining and Measuring Human Well-Being." In *Global Environmental Change: Handbook of Global Environmental Pollution*, edited by B. Freedman, 833–851. Dordrecht: Springer.

Cowen, R. 2014. "Ways of Knowing, Outcomes and 'Comparative Education': Be Careful What you Pray for." *Comparative Education* 50 (3): 282–301.

Crossley, M. 2014. "Global League Tables, Big Data and the International Transfer of Educational Research Modalities." *Comparative Education* 50 (1): 15–26.

Davis, K., B. Kingsbury, and S. Merry. 2012. "Indicators as a Technology of Global Governance." *Law & Society Review* 46 (1): 71–104.

DeJaeghere, J. 2015. "Reframing Gender and Education for the Post-2015 Agenda." In *Routledge Handbook of International Education and Development*, edited by S. McGrath and Q. Gu, 63–77. Abingdon: Routledge.

Esquivel, V. 2016. "Power and the Sustainable Development Goals: A Feminist Analysis." *Gender & Development* 24 (1): 9–23.

Fredman, S., J. Kuosmanen, and M. Campbell. 2016. "Transformative Equality: Making the Sustainable Development Goals Work for Women." *Ethics & International Affairs* 30 (2): 177–187.

Fukuda-Parr, S., T. Lawson-Remer, and S. Randolph. 2015. *Fulfilling Social and Economic Rights*. New York: Oxford University Press.

Fukuda-Parr, S., A. Yamin, and J. Greenstein. 2014. "The Power of Numbers: A Critical Review of Millennium Development Goal Targets for Human Development and Human Rights." *Journal of Human Development and Capabilities* 15 (2–3): 105–117.

Goldstein, H., and G. Moss. 2014. "Knowledge and Numbers in Education." *Comparative Education* 50 (3): 259–265.

Gore, C. 2015. "The Post-2015 Moment: Towards Sustainable Development Goals and A New Global Development Paradigm." *Journal of International Development* 27 (6): 717–732.
Gorur, R. 2014. "Towards A Sociology of Measurement in Education Policy." *European Educational Research Journal* 13 (1): 58–72.
Henderson, E. 2015. *Gender Pedagogy*. London: Palgrave.
Kabeer, N. 1999. "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment." *Development and change* 30 (3): 435–464.
Kertzer, D., and D. Arel, ed. 2002. *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*. Cambridge: Cambridge University Press.
King, E., and A. Hill, eds. 1993. *Women's Education in Developing Countries: Barriers, Benefits, and Policies*. Washington: World Bank Publications.
Lawn, M., ed. 2013. *The Rise of Data in Education Systems: Collection, Visualization and use*. Oxford: Symposium Books.
Loots, S., and M. Walker. 2015. "Shaping A Gender Equality Policy in Higher Education: Which Human Capabilities Matter?" *Gender and Education* 27 (4): 361–375.
Marx, A., B. Rihoux, and C. Ragin. 2014. "The Origins, Development, and Application of Qualitative Comparative Analysis: the First 25 Years." *European Political Science Review* 6 (01): 115–142.
Meyer, H. D., and A. Benavot, eds. 2013. *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Oxford: Symposium Books.
Mishra, P. 2016. "Straitjacketing education: An investigation of heteronormativity in the 2030 Education Agenda." Unpublished MA diss., University College London, Institute of Education, London.
Miske, S., M. Meagher, and J. DeJaeghere. 2010. "Gender Mainstreaming in Education at the Level of Field Operations: the Case of CARE USA's Indicator Framework." *Compare* 40 (4): 441–458.
Morris, P. 2015. "Comparative Education, PISA, Politics and Educational Reform: A Cautionary Note." *Compare* 45 (3): 470–474.
Mundy, K., and C. Manion. 2015. "The Education for All Initiative: History and Prospects Post-2015." In *Education and International Development: Practice, Policy and Research*, edited by T. McCowan and E. Unterhalter, 49–68. London: Bloomsbury.
Nussbaum, M. 2000. *Women and Human Development*. Cambridge: Cambridge University Press.
Nussbaum, M. 2011. *Creating Capabilities. The Human Development Approach*. Cambridge: Harvard University Press.
OECD. 2015. *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Paris: OECD.
Oya, C. 2015. "Who Counts? Challenges and Biases in Defining "Households" in Research on Poverty." *Journal of Development Effectiveness* 7 (3): 336–345.
Parkes, J., ed. 2015. *Gender Violence in Poverty Contexts: the Educational Challenge*. Abingdon: Routledge.
Parkes, J., J. Heslop, F. Johnson Ross, R. Westerveld, and E. Unterhalter. 2016. *A Rigorous Review of Global Research Evidence on Policy and Practice on School Related Gender Based Violence*. New York: UNICEF.
Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. London: Sage.
Pogge, T. and M. Sengupta. 2016. "Assessing the Sustainable Development Goals From a Human Rights Perspective." *Journal of International and Comparative Social Policy* 32 (2): 83–97.
Punton, M., I. Vogel, and R. Lloyd. 2016. "Reflections from a Realist Evaluation in Progress: Scaling Ladders and Stitching Theory." CDI Practice Paper 18, Brighton: Institute of Development Studies.
Randall, S. and E. Coast. 2015. "Poverty in African Households: the Limits of Survey and Census Representations." *The Journal of Development Studies* 51 (2): 162–177.
Robeyns, I. 2003. "Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities." *Feminist Economics* 9 (2–3): 61–92.
Robeyns, I. 2005. "Selecting Capabilities for Quality of Life Measurement." *Social indicators research* 74 (1): 191–215.
Robeyns, I. 2016. "Capabilitarianism." *Journal of Human Development and Capabilities* 17 (3): 397–414.
Rottenburg, R., S. Merry, S. Park, and J. Mugler. 2015. *The World of Indicators: the Making of Government Knowledge Through Quantification*. Cambridge: Cambridge University Press.

Sayed, Y. 2013. "The Post-2015 Education and Development Agenda." *Compare* 43 (6): 783–846.
Scheyvens, R., G. Banks, and E. Hughes. 2016. "The Private Sector and the SDGs: The Need to Move Beyond 'Business as Usual'." *Sustainable Development*. doi:10.1002/sd.1623
Sen, A. 1999. *Development as Freedom*. Oxford: Oxford Univeristy Press.
Sen, A. 2004. "Capabilities, Lists, and Public Reason: Continuing the Conversation." *Feminist Economics* 10 (3): 77–80.
Sen, A. 2005. " Human Rights and Capabilities." *Journal of Human Development* 6 (2): 151–166.
Sen, G., and A. Mukherjee. 2014. "No Empowerment Without Rights, No Rights Without Politics: Gender-Equality, MDGs and the Post-2015 Development Agenda." *Journal of Human Development and Capabilities* 15 (2–3): 188–202.
Sénit, C., F. Biermann, and A. Kalfagianni. 2016. "The Representativeness of Global Deliberation: A Critical Assessment of Civil Society Consultations for Sustainable Development." *Global Policy*. doi:10.1111/1758–5899.12371
Simon, P., and V. Piché. 2012. "Accounting for Ethnic and Racial Diversity: the Challenge of Enumeration." *Ethnic and Racial Studies* 35 (8): 1357–1365.
de Sousa Santos, B. 2015. *Epistemologies of the South: Justice Against Epistemicide*. Abingdon: Routledge.
Tabbush, C. 2014. "Posing Local Challenges in Global Debates: Comments on Poverty Eradication, Human Rights and Gender Justice in Latin America." Contribution to consultation on global justice, poverty and inequality in the post-2015 development agenda, 28–29 April. Paris: UNESCO.
United Nations. 2015. *The Sustainable Development Goals*. New York: United Nations.
UNDP. 2015. *Human Development Report. Work for Human Development*. New York: United Nations Development Programme.
UN ECOSOC. (2016). "Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators." UN Economic and Social Council Statistical Commission 47th Session, E/CN.3/2016/2/Rev.1.
UNESCO. 2016. *Education for People and Planet: Creating Sustainable Futures for all*. Paris: UNESCO.
Unterhalter, E. 2009. "Translations and Transversal Dialogues: an Examination of Mobilities Associated with Gender, Education and Global Poverty Reduction." *Comparative Education* 45 (3): 329–345.
Unterhalter, E. 2014a. "Measuring Education for the Millennium Development Goals: Reflections on Targets, Indicators, and a Post-2015 Framework." *Journal of Human Development and Capabilities* 15 (12): 176–187.
Unterhalter, E. 2014b. "Thinking About Gender in Comparative Education." *Comparative Education* 50 (1): 112–126.
Unterhalter, E. 2015. "Measuring gender inequality and equality in education." Concept note for UNGEI workshop, London. https://www.researchgate.net/publication/283084293_Measuring_Gender_inequality_and_equality_in_education.
Vandemoortele, J. 2011. "The MDG Story: Intention Denied." *Development and Change* 42 (1): 1–21.
Waage, J., J. Banerji, O. Campbell, E. Chirwa, G. Collender, V. Dieltiens, A. Dorward et al. 2010. "The Millennium Development Goals: A Cross Sectoral Analysis and Principles for Goal Setting Post 2015." *The Lancet* 376 (9745): 991–1023.
Walby, S. 2011. *The Future of Feminism*. Cambridge: Polity.