

Measuring what we Value or Valuing what we Measure? Globalization, Accountability and the Question of Educational Purpose

¿Medir lo que valoramos o valorar lo que medimos? Globalización, responsabilidad y la noción de propósito de la educación

Gert Biesta

University of Luxembourg

Abstract

One of the significant dimensions of the impact of globalisation on educational policy and practice has been the rise of large-scale comparative measurements of the performance of national educational systems. In this paper I argue that rather than that this development has supported and promoted discussions about good education, they have actually replaced normative questions about desirable educational orientations and achievements with technical questions about the effective production of a particular (and often narrow) set of educational outcomes. On the one hand I aim to contribute to the understanding of why this has been the case, and here I particularly highlight a shift in thinking about accountability in education from a substantive and democratic approach to a technical-managerial approach. On the other hand I provide parameters for a more explicit engagement with questions about what is educationally desirable, *not* in order to specify once and for all what good education is or should be, but in order to facilitate a more sophisticated, nuanced and deliberate discussion about what the parameters of good education in schools, colleges, universities and other educational settings and institutions might be.

Keywords: globalisation, measurement, league tables, accountability, school effectiveness, democracy, good education

Post to:

Gert Biesta
Faculty of Language and Literature, Humanities, Arts and Education
University of Luxembourg, Campus Walferdange
Route de Diekirch, BP2, L-7220 Walferdange, Luxembourg
Email: gert.biesta@uni.lu

© 2014 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN: 0719-0409 DDI: 203.262, Santiago, Chile
doi:10.7764/PEL.51.1.2014.5

Resumen

Una de las dimensiones más importantes del impacto de la globalización en las políticas y prácticas de la educación ha sido el aumento de las mediciones comparativas a gran escala del desempeño de los sistemas de educación nacionales. En este artículo, mi argumento es que más que este desarrollo haya respaldado y promovido un debate respecto de la buena enseñanza, estas mediciones han reemplazado las preguntas normativas sobre las orientaciones y logros educacionales deseables con preguntas técnicas sobre la producción eficaz de un conjunto particular (y con frecuencia delimitado) de resultados educacionales. Por un lado, apunto a contribuir a la comprensión de por qué se ha dado esta situación, y aquí destaco en particular un cambio desde un enfoque sustancial y democrático a un enfoque técnico y administrativo respecto de cómo se considera la rendición de cuentas en la educación. Por otro lado, brindo parámetros para una discusión más explícita que se centre en las preguntas acerca de qué se busca obtener en la educación, *no* con el fin de especificar de una vez por todas lo que es o debería ser la buena enseñanza, sino más bien para facilitar un debate más sofisticado, variado y reflexivo acerca de cuáles podrían ser los parámetros de una buena educación en escuelas, colegios, universidades y en otros entornos e instituciones de educación.

Palabras clave: globalización, medición, clasificaciones, rendición de cuentas, eficacia escolar, democracia, buena enseñanza

The construction of the educational globe

While a substantial part of the literature on globalisation focuses on the economic dimension, that is, on the ongoing reduction of barriers between national borders in order to facilitate the free flow of goods, capital, services and labour, globalisation also occurs in other domains and with respect to other dimensions, such as politics and policy, culture and technology (Michie, 2011). In such areas globalisation tends to result in the convergence of processes, practices and ideas, and thus in an increased *uniformity* of ways of doing, being and thinking. Although educational systems tend to be strongly rooted in national histories and identities —both as the outcome *of* and as a factor *in* nation building (Green, 1990)— the last few decades have witnessed a strong rise in the globalisation of educational policy and practice (e.g. Rizvi & Lingard, 2009). While educational globalisation may partly be the result of relatively ‘spontaneous’ processes, and while there is also clear evidence of the effects of ‘policy borrowing’ (Phillips, 2005; Phillips & Ochs, 2004), the increasing globalisation of educational discourse and practice is also the outcome of very concrete and specific interventions by supra-national bodies and organisations such as the World Bank, UNESCO and the OECD.

Some of such interventions are clearly *substantive*. Both the World Bank and UNESCO, for example, have clear, albeit differing views about the promotion of educational development around the world. But another important driver of educational globalisation stems from much more *formal* interventions in the field in the form of the large-scale comparative measurement of educational ‘outcomes’ through such systems as the *Trends in International Mathematics and Science Study* (TIMSS), the *Progress in International Reading Literacy Study* (PIRLS) and, most notoriously, OECD’s *Program for International Student Assessment* (PISA) (Martens, Nagel, Windzio, & Weymann, 2010). While one could argue that such systems in themselves merely measure what is already ‘there,’ their actual impact goes much further. This is because of the fact that many countries tend to adjust their policies and practices in response to and in anticipation of the outcomes of such measurements in order to gain a higher place in the league tables that such systems produce. It is, therefore, not just the impact of such measurements in themselves, but more importantly the response to such systems of comparative measurement that brings about adjustments of national educational policy and practices toward the alleged ‘standard’ implied in such systems (Waldow, 2009). It is in precisely this way that such systems contribute to the ongoing standardisation, harmonisation and unification of the ‘educational globe’ (Tröhler, 2010).

In this contribution I wish to raise a number of critical questions about these developments, particularly with regard to their impact on ideas about quality and standards in education or, in the language I prefer (Biesta, 2010), with regard to their impact on views about *good* education. The thesis I will put forward

is that rather than large-scale comparative measurement of educational outcomes has supported and promoted discussions about good education, it has actually replaced normative questions about desirable educational orientations and achievements with technical questions about the effective production of a particular (and often narrow) set of educational outcomes. On the one hand I aim to contribute to the understanding of why this has been the case—and here I will particularly highlight a shift in thinking about accountability in education from a substantive and democratic approach to a technical-managerial approach. On the other hand I aim to provide parameters for a more explicit engagement with questions about what is educationally desirable, *not* in order to specify once and for all what good education is or should be, but in order to facilitate a more sophisticated, nuanced and deliberate discussion about what the parameters of good education in schools, colleges, universities and other educational settings and institutions might be. An ongoing discussion about what our educational efforts should aim to achieve is, so I will argue, of crucial importance if we do not wish to hand over responsibility for our educational processes and practices to abstract systems of measurement, but aim to keep democratic control over our educational endeavours and over the ways in which we assess their quality.

The rise of a culture of measurement

Over the past decades there has been a remarkable rise of interest in the measurement of education or, in the words of those involved, in the measurement of educational ‘outcomes.’ This has not been confined to the Western world, but is rapidly becoming a global phenomenon, not least through the involvement of organisations such as the OECD, UNESCO and the World Bank. The most prominent manifestations of the rise of a global culture of measurement can be found in the international comparative studies mentioned above that result in league tables which are assumed to indicate who is better and who is best. Such league tables are intended to provide information about how national education systems perform compared with those of other countries. They are therefore generally competitive in their outlook as there can ultimately be only one ‘number one.’

Findings from such studies are utilised by national governments to inform education policy, often under the banner of ‘raising standards,’ ‘creating excellence,’ or ‘keeping ahead in the global economy’. League tables are also produced at national level with the aim of providing information about the relative performance of individual schools or school districts. Such league tables have a complicated rationale, combining elements of accountability, selectivity and control with a social justice argument which says that everyone should have access to education of the same quality (Webb, 2011). The data used for producing such league tables are also used to identify so-called ‘failing schools’ and, in some cases, ‘failing teachers’ within schools, and in this respect clearly contribute to a climate of ‘naming and shaming’ and of blaming individuals for what are often the outcomes of the complex interaction of a wide range of structural issues that lie well beyond the control of individual teachers or schools (e.g. Granger, 2008; Hess, 2006; Kumashiro, 2012; Nicolaidou & Ainscow, 2005; Tomlinson, 1997).

The rise of the measurement culture in education has had a profound impact on educational practice from the highest levels of educational policy down to the everyday activities of schools and teachers. To some extent this impact has been beneficial as it has made it possible to base discussions on factual information rather than just on assumptions or opinions about what might be the case. The problem, however, is that the abundance of information about educational outcomes has given the impression that decisions about the direction of educational policy and the shape and form of educational practice can *solely* be based upon factual information. There are, however, at least two problems with this way of thinking.

The first is that while it is always advisable to use factual information when making decisions about what ought to be done, what ought to be done can never be logically derived from what is the case. This problem, which in the philosophical literature is known as the ‘is-ought problem’, implies that when we engage in decisions about the direction of education we always and necessarily have to make *value* judgments—judgments about what is educationally *desirable*. This implies that if we wish to say something about the direction of education we always need to complement factual information with views about what is considered to be desirable. We need, in other words, to *evaluate* data and evidence and for this, as has been known for a long time in the field of educational evaluation, we need to engage with *values* (e.g. Henry, 2002; House & Howe, 1999; Schwandt & Dahler-Larsen, 2006).

The second problem, which is related to the first and in a sense is its methodological corollary, is the problem of the validity of our measurements. More than just the question of the *technical validity*—that is, the question of whether we are measuring what we intend to measure—the problem lies in what I suggest to refer to as the *normative validity* of our measurements. This has to do with the question of whether we are indeed measuring what we value, or whether we are just measuring what can easily be measured so that we end up in a situation where we value what we can measure or what has been measured. The rise of a culture of performativity in education—a culture in which means become ends in themselves so that targets and indicators of quality become mistaken for quality itself—has been one of the main drivers of an approach in which normative validity (where we measure what we value) is being replaced by technical validity (where we are supposed to value what is measurable) (e.g. Ball, 2003; Usher, 2006).

In order to understand the rise of a culture of measurement in education, its background and its impact, we need to turn our attention to the context in which these changes have occurred. This brings me to the question of accountability.

The culture of measurement in the context of accountability

In a recent analysis of the PISA phenomenon, Stefan Hopmann asks the simple but important question: “What is different with PISA?” (Hopmann, 2008, p. 417). The reason for asking this question is that PISA is not the first attempt at compiling comparative international data about educational performance. The IEA, the International Association for the Evaluation of Educational Achievement—the organisation behind such studies as TIMSS, PIRLS and, more recently ICCS, TIMSS Advanced and TEDS-M, to add some more acronyms to the mix (see www.iea.nl)—has, after all, been around since the late 1950s. What, then, explains the ‘success’, as Hopmann calls it—or the ‘impact’ as I would prefer to say—of PISA? Hopmann argues that part of the answer to this question has to be sought not within PISA itself but within the changing context, the changing ‘social environment’ in which PISA operates (Hopmann, 2008, p. 418). The factor singled out by Hopmann to explain the impact of studies like PISA is *accountability* and, more specifically, the rise in the influence of a particular approach to accountability as a way in which societies—particularly in the Western world—deal with welfare problems like security, health, and education (Hopmann, 2008, p. 418).

That we need data in order to be accountable is, in itself, not really a problematic contention. But the question here is not only about the kind of data we might need; the question is also about the kind of accountability we are aiming for. With regard to the latter question Charlton (1999, 2002) has made a helpful distinction between two largely distinct conceptions of accountability: a *technical-managerial* conception and a looser, more general conception, which, as I will argue below, can be characterised as a *professional* or *democratic* conception of accountability. In general discourse, accountability has to do with responsibility and carries connotations of ‘being answerable-to’. The technical-managerial meaning refers more narrowly to the duty to present auditable accounts. Originally accountability only referred to *financial* documentation. The current *managerial* use of accountability is a direct extension of this financial usage in that an accountable organisation is seen as one that has the duty to present auditable accounts of *all* of its activities.

The link between the technical-managerial interpretation of accountability and the professional-democratic interpretation is weak. Charlton argues that “[only] insofar as it is legitimate to assume that the provision of auditable documentation is synonymous with responsible behaviour” is there any overlap between the two meanings of accountability (Charlton, 2002, p. 18). Yet, the rhetoric of accountability operates precisely on the basis of a ‘quick switch’ (Charlton) between the two meanings, making it difficult to see an argument *against* accountability as anything other than a plea *for* irresponsible action.

Charlton not only makes a helpful conceptual distinction between the two interpretations of accountability; his account also shows that the history of *managerial* use of the idea of accountability lies in a strictly financial context in which the purpose of auditing is “to detect and deter incompetence and dishonesty in the handling of money” (Charlton, 2002, p. 24). He argues that the logic of financial auditing has simply been *transposed* to the managerial context, without much consideration for the question to what extent this logic is appropriate for managerial purposes. Rather than adapting the principles of

the auditing process to the specifics and requirements of a different context, Charlton demonstrates that the culture of accountability has led to a situation in which practices have had to adapt to the principles of the auditing process (Power, 1994, 1997). This is based on the following train of thought: “Transparent organizations are auditable, and auditable organizations are manageable —and *vice versa*. Therefore, organizations *must be made auditable*” (Charlton, 2002, p. 22).

Although Charlton seems to suggest that the two meanings of accountability currently coexist, it could be argued that the tradition which sees accountability as a system of (mutual) responsibility rather than as a system of governance and control was the dominant tradition before the rise of the technical-managerial approach. There is clear evidence for this in education, where, as Poulson (1996, 1998) has shown, discussions about accountability in the late 1970s and early 1980s were strongly focused on a *professional* interpretation of accountability, while there were also attempts in education to articulate a *democratic* approach to accountability, arguing that making schools accountable to parents, students and the wider citizenry would support the democratisation of education (Davis & White, 2001; Epstein, 1993). But this, so it appears, is no longer the case.

The reconfiguration of the relationship between the state and its citizens

The shift from professional and democratic notions of accountability to the current hegemony of the technical-managerial approach should be understood against the background of wider changes in society. Many authors have argued that the rise of the technical-managerial approach to accountability is connected to *ideological* transformations (most notably the rise of neo-liberalism and neo-conservatism) and *economic* changes (most importantly the oil-crisis and the economic slowdown in the mid 1970s, and the subsequent rise of global capitalism) which, together, have resulted in the decline —if not demolition— of the welfare state and the rise —if not hegemony— of the neo-liberal/global capitalist logic of the market (see also Hopmann, 2008, p. 423). One of the most significant changes that has been brought about as a result of these developments has been the reconfiguration of the relationship between the state and its citizens. This relationship has become less a political relationship —that is, a relationship between government and citizens who, together are concerned about the common good— and has increasingly become an economic relationship where, in the first stage, the state is the provider and the taxpayer the consumer of public services and where, in the second stage, the state leaves the provision of such services to private providers and becomes a regulator of the ‘public service market’ (which I put in quotation marks in order to highlight that as soon as public services enter the market place they lose their public character —see also below).

The reconfiguration of the relationship between the state and its citizens should not be understood as simply a different way of relating. The new relationship has fundamentally changed the role and identity of the two parties and the terms on which they relate. Not only can it be argued that the relationship between the state and its citizens has been de-politicised. One could even argue that the sphere of the political itself has been eroded (Biesta, 2005; Marquand, 2004). Crucially, the language that is used is an economic language which positions the government as provider and the citizen as consumer (Biesta, 2004, 2006). *Choice* has become the keyword in this discourse. Yet ‘choice’ is about the behaviour of consumers in a market where their aim is to satisfy their needs, and should not be conflated with democracy, which is about public deliberation and contestation about the common good and the just and equitable (re)distribution of public resources.

According to the logic of the market, the relationship between the state and its citizens is no longer a *substantial* relationship, but has turned into a strictly *formal* relationship. This reconfiguration is closely connected to the rise of quality assurance. Indeed, current quality assurance practices typically concentrate “upon *systems* and *processes* rather than outcomes” (Charlton, 2002, p. 20), which is not to suggest that outcomes are not deemed relevant, but that they are seen as outside the ‘reach’ of the question of quality; that is, the question as to what desirable ‘outcomes’ are and who should have a say in defining them. This is why, for example, the constant emphasis of the British government on ‘raising standards’ in education and other public services is rather vacuous since it lacks proper (democratic) discussion about which standards or ‘outcomes’ are most desirable. The same problem underlies much of the earlier research of the “school effectiveness and improvement industry” (Gewirtz, 2002, p. 15), since these studies mainly focused on the effectiveness and efficiency of processes, without raising the far more difficult normative

and political question about the desirability of what such processes should result in (Townsend, 2007 includes an overview of more recent developments in this field).

From political to economic relationships

What all this shows is that the rise of the managerial approach to accountability is not an isolated phenomenon, but that it is part of a larger transformation of society in which political relationships and the sphere of the political itself seem to have been replaced by economic relationships. The ground for the current mode of accountability seems to be an economic one, in that the right to accountability that the government claims seems to arise from the financial investment it makes in public services like education. Although at first sight there seem to be opportunities for a more democratic ‘face’ of accountability, that is, in the relationship between parents and students as ‘consumers’ of education and schools as ‘providers,’ one of the problems is that there is no direct relationship of accountability between these parties, but only an indirect one. The only role parents and students can play is that of consumers of educational provision, but there is no opportunity to participate in any public, democratic discourse about education. Onara O’Neill describes the predicament as follows:

In theory the new culture of accountability and audit makes professionals and institutions more accountable *to the public*. This is supposedly done by publishing targets and levels of attainment in league tables, and by establishing complaint procedures by which members of the public can seek redress for any professional or institutional failures. But underlying this ostensible aim of accountability *to the public* the real requirements are for accountability *to regulators, to departments of government, to funders, to legal standards*. The new forms of accountability impose forms of *central control*—quite often indeed a *range of different and mutually inconsistent* forms of central control (O’Neill, 2002, p. 4).

The problem is that while many would want the culture of accountability to do the former (i.e., to be accountable to the public), it actually does the latter (i.e., being accountable to regulators) and thereby takes the real stakeholders out of the ‘accountability loop.’ In this respect the current technical-managerial approach to accountability produces economic relationships between people and makes democratic relationships difficult, if not impossible.

The impact of this on day-to-day practice in schools and other institutions is that institutions seem to adapt themselves to the requirements of accountability and auditing, rather than the other way around. To quote O’Neill once more:

In theory again the new culture of accountability and audit makes professionals and institutions more accountable *for good performance*. This is manifest in the rhetoric of improvement and raising standards, of efficiency gains and best practice, of respect for patients and pupils and employees. But beneath this admirable rhetoric the real focus is on performance indicators chosen for ease of measurement and control rather than because they measure accurately what the quality of performance is (O’Neill, 2002, pp. 4-5).

O’Neill points out that the incentives of the culture of accountability are by no means unreal. Yet what they seem to elicit is behaviour that suits the accountability system —behaviour that suits the inspectors and those responsible for quality assurance— rather than acting as an incentive for professional and responsible action. Ironically, this can easily result in a situation that is detrimental for the ‘consumers’ of public services. If, for example, schools are rewarded for high exam scores, they will increasingly only try to attract ‘motivated’ parents and ‘able’ children and will try to keep ‘difficult students’ out. Ultimately, this results in a situation where it is no longer the question of what schools can do for their students, but what students can do for their school (Apple, 2000, p. 235; Hopmann, 2008, pp. 443-444).

Measuring what we value?

The foregoing analysis of the rise of a particular ‘regime’ of accountability in the wake of the decline of the welfare state and the rise of the market orientation of neo-liberalism helps to explain what is different about the context in which contemporary large-scale measurement of education takes place. More specifically it helps to explain the remarkable impact of the culture of measurement, as measurement, and particularly comparative measurement, is the ‘fuel’ for the technical-managerial approach to accountability. After all, technical-managerial accountability is only possible if there is ongoing information about the performance of the system. Although the abundance of data about the relative performance of public

services may give the impression of transparency and openness, the problem is that many of the real ‘stakeholders’ have been taken out of the accountability loop. In this regard the technical-managerial approach to accountability is disempowering rather than empowering, also because the rhetoric of choice is often confined to choice from a set menu rather than interested parties being able to have a democratic say in what is on the menu in the first place. This is not to suggest, too simply, that in the past democracy was fully realised while in the present it has completely disappeared, but it does highlight that the logic of choice is fundamentally opposite to the logic of democracy.

The foregoing analysis not only helps us to understand the impact, influence and success of the culture of measurement under the condition of technical-managerial accountability. It suggests at the very same time that this not the only way in which questions about educational quality and educational accountability can be addressed. It suggests, in other words, that there is a real choice in that accountability does not necessarily have to be understood in the technical-managerial sense, but that there is a *democratic* alternative. This is the alternative in which stakeholders are not taken out of the accountability loop, but play a central role in it, particularly with regard to decisions about what should be on offer instead of only being allowed to choose from an offer defined by others. And it is an alternative where the focus is not, as O’Neill has put it, on performance indicators “chosen for ease of measurement and control” (O’Neill, 2002, p. 5) —the situation in which we are ‘valuing what we (can) measure— but where the focus is on the measurement of what is valued. Such an approach to accountability therefore requires a different kind of measurement, one that is not just generating comparative data in order to indicate who is better and who is best —thus generally contributing to a focus on competition rather than cooperation— but one where there is a genuine concern for quality: not the quality of processes, but the quality of what such processes are supposed to bring about. This is not simply to ask that we pay attention to ‘outcomes’ —as contemporary accountability and measurement regimes in a sense pay a lot of attention to ‘outcomes’— but to ask the deeper normative and political question about which ‘outcomes’ are considered desirable, not just in terms of individual preference (the logic of choice), but in terms of what is considered collectively desirable (the logic of democracy, which is a logic that precisely may limit or transform the preferences of some for the sake of the greater common good). When measurement operates in function of such a more empowering approach to accountability it therefore needs to engage explicitly with questions of purpose —which brings me to the second theme of this paper, that is, the question of educational purpose and of the ‘goodness’ of education more generally.

The question of purpose: What is education for?

Earlier in this paper I made a distinction between *technical validity* and *normative validity*. Whereas technical validity has to do with the question of whether we are indeed measuring what we are supposed to measure, normative validity —in my definition— has to do with the question of whether we are indeed measuring (or at least trying to measure) what we value. In the domain of education this immediately brings us to the question of purpose —the question as to what education is *for*— since it is only in function of particular ideas about what we aim to achieve in and through our educational activities and efforts that measurement of educational ‘outcomes’ can in any way be meaningful and thus can carry normative validity. In the current culture of measurement this question seems relatively absent. Or perhaps it is more accurate to say that such studies seem to rely on an unquestioned ‘common sense’ view of what education is for, focusing on a particular and generally narrow set of ‘outcomes’ (for a discussion of these problems with regard to PISA see Hopmann, 2008, pp. 438-440). This can easily result in a situation where what is measured and measurable becomes that which is valued, particularly if it is politically ‘convenient’, so to speak, to appeal to such a particular, narrow —and perhaps we can even call this populist— definition of what good education is. (Note that PISA only focuses on a very narrow set of curricular subjects —science, reading and mathematics— and therefore represents a very narrow and specific conception of what is supposed to ‘matter’ educationally.) This, in turn, can lead to a situation in which measurement and policy feed into each other and mutually reinforce each other without ever asking whether the direction of policy and the direction of the measurement that feeds into the policy are in themselves desirable (for an example of such a problematic relationship see Biesta, 2009b).

But the issue is not just one of common sense or populist views about what the purposes of education are. For some reason —which I actually find quite difficult to put my finger on— there is a more general lack of attention for and clarity about the question as to what education is for, at least, that

is, in the English speaking world. This is not to say that there is an absence of statements about what good education is, but these statements often remain empty. Let me give two examples of this. One can be found in the claim that the purpose of education is that students learn, sometimes articulated in more precise terms, such as, that students engage in active learning or in collaborative learning. While ‘learning’ as an aim for education sounds good, it actually means very little if we do not specify *what* students are supposed to learn and, even more importantly, *why* they should learn this. While saying that the purpose of education is that students learn may seem to express a view about what education is for, it actually says very little (Biesta, 2013).

There is a similar problem with a concept that has actually had a huge impact on the development of a culture of measurement in education, which is the notion of effectiveness. One might argue that the ambition to improve the effectiveness of education, be it at the level of the school system, individual schools, year groups, or individual teachers and their teaching, states a clear purpose for education. Moreover it does so through the use of value language, as it cannot be disputed that ‘effectiveness’ is an evaluative term. The problem here, however, is that effectiveness is an *instrumental* value, a value that expresses something about the ability of certain processes to bring about certain outcomes. But the idea of effectiveness is neutral with regard to the desirability of the outcomes (as testified by the fact that there can be such a thing as effective torture, for example). This is why the case for effective education is not enough, and in a sense is actually misleading. There is always the question ‘Effective for what?’, and given that what is effective for one particular student or group of students may not necessarily be effective for other students, there is also the additional question ‘Effective for whom?’ (Bogotch, Mirón, & Biesta, 2007).

Functions of education and domains of educational purpose

Both ‘learning’ and ‘effectiveness’ are therefore in themselves inadequate concepts for saying anything about the question of educational purpose, that is, about those outcomes of education that are considered to be desirable. Although any answer to this question is, of course, contentious and—in the positive sense of the word—ideological, I have found it helpful in discussions about the purpose or purposes of education (and particularly school education) to highlight the fact that education functions in relation to a number of different dimensions, that such dimensions are not necessarily in synergy with each other, and that different school concepts, educational philosophies and even pedagogies articulate a different position in relation to these dimensions. The most important point here is to acknowledge that ‘education’ is a composite concept. This is reflected in the fact that educational practices are not mono-functional, but in reality perform a number of different functions (and often do so at the very same time). With regard to this I have suggested distinguishing between three functions of education, to which I refer as qualification, socialisation and subjectification (for reasons of space I introduce these concepts here briefly in order to indicate what it would entail to engage with questions of purpose in education. (For a much more detailed discussion of these ideas see Biesta, 2009a, 2010, 2014).

Qualification has to do with the ways in which education contributes to the acquisition of knowledge, skills and dispositions that qualify us for doing something—a doing which can range from the very specific (such as training for a particular job) to the very general (such as in the case of liberal education). *Socialisation* has to do with the ways in which, through educational processes and practices, individuals become part of existing socio-cultural, political and moral orders. Schools partly engage in socialization deliberately, for example, in the form of values education, character education or citizenship education, or in relation to professional socialisation. Socialisation also happens in less visible ways, as has been made clear in the literature on the hidden curriculum and the role of education in the reproduction of social inequality. Whereas some would argue that education should only focus on qualification, and others defend that education has an important role to play in the socialization of children and young people, there is a third function of education which is different from both qualification and socialization. This function has to do with the ways in which education contributes to the individuation or, as I prefer to call it, the *subjectification* of children and young people. The subjectification function might perhaps best be understood as the opposite of the socialization function. It is *not* about the insertion of ‘newcomers’ into existing orders, but about ways of being that hint at independence from such orders; ways of being in which the individual is not simply a ‘specimen’ of a more encompassing order.

The point of making these distinctions is not only to argue that education can potentially impact on quite different dimensions (which would be to see them as three different functions of education). The same dimensions also play an important role in the *justification* of educational processes and practices, and thus specify different views—or different dimensions of views—about what education is for (which means that they can also be seen as three different domains of educational purpose, i.e. three domains within which we can and should engage with the question what education is *for*). Although there are examples of justifications of (school) education that focus on only one of these dimensions—the most prominent case being the view that school education should only operate in the domain of qualification—most justifications of education contain a particular mix of these dimensions, and one of the important practical questions for educators is how the three dimensions can be kept ‘in balance,’ so to speak, as an emphasis on one dimension may have a negative impact on what can be achieved in another dimension. For this reason I prefer to depict the three functions/purposes of education in a Venn diagram of three partly overlapping circles, where the overlap highlights both the potential for synergy—that is, that work in the domain of qualification can support meaningful change in, say, the domain of subjectification—and for conflict, that is, where work in the domain of, say, socialisation, goes against what we aim to achieve in the domain of subjectification. The idea of ‘balance,’ to put it differently, is not to be thought of in quantitative terms, but rather in a qualitative way. Looking at it in this manner also helps to see the distortion of a meaningful balance that occurs when measurement only focuses on one dimension (in most cases this being the dimension of qualification).

My reason for highlighting the fact that education can perform different functions and serve different purposes is not to get into a discussion about what education is or should be for—that is precisely what should be a topic of ongoing concern, deliberation and contestation in democratic societies—but to emphasise that engagement with the question of purpose in educational measurement requires a multi-dimensional approach, or at least an approach that is sensitive to the range of different ideas about what is educationally desirable. It is only when we start from here that we can move towards a situation in which we measure what is valued, rather than ending up in a position where we value what is or what can be measured.

Discussion and conclusions

In this paper I started from the observation that the current educational climate in many countries around the world is characterised by an abundance of measurement, particularly large-scale international comparative measurement. I have referred to this as a 'culture of measurement' and have highlighted some aspects of the problematic impact of this on educational practice, particularly in relation to processes of educational globalisation. One problem has to do with the way in which the culture of measurement contributes to the constant surveillance and control of educational processes and practices. Another problem has to do with the fact that it contributes to a culture of competition and perhaps even to a culture of fear, where the ambition to stay ahead of everyone else is linked to a fear of being left behind. Also, the current culture of measurement operates on a global scale and, as a result, contributes strongly to a convergence of national systems of education towards the particular definition of good education promoted by global measurement systems (for an analysis of similar processes with regard to higher education see Biesta, 2011). I have suggested that the impact of the current culture of measurement should be understood against the background of a very particular regime of accountability, a regime to which I have referred as technical-managerial accountability. The rise of this regime should itself be understood against the background of the decline of the welfare state and the rise of neo-liberal forms of government and governance.

I have also shown, however, that there are two readings of accountability. There is not only the technical-managerial interpretation in which accountability is basically a system of central control with disempowering and anti-democratic effects. There is also a democratic reading of accountability where stakeholders are within the accountability loop rather than being outside it, and where the focus is not on the question of how everyone is positioned relative to everyone else, but where the focus is on questions of substance such as, in the case of education, the question as to what good education is. I have shown that the latter question is a composite question, since education always works in a number of different dimensions and often does so at the very same time. This not only means that justifications of education and articulations about what is educational desirable will have to be multi-dimensional. It also implies that any attempt to measure or assess the achievements of education needs to take the multi-dimensional character of education into consideration. This not only requires explicit attention to these dimensions, but also raises questions about what the most appropriate ways to assess achievements and outcomes in these different areas are. It also requires, in other words, the development of 'dimension-appropriate' ways of evaluation and assessment.

It is for these reasons that I have argued for the need to engage with the question of purpose in the measurement, assessment and evaluation of education, as it is only when this is taken into consideration that we can move on from a position in which we are valuing what is or can be measured to a situation in which we allow our judgements about education to be informed by the measurement of what we value. The emphasis here is not only on 'value', that is, measuring what we *value*; the emphasis also needs to be on the 'we', that is, measuring what *we* value. For measurement to be a positive force in empowering and in democratic forms of accountability it is, after all, important that we do not focus on what is valued in the abstract sense, but give a voice to all stakeholders and interested parties in articulating what it is that is considered to be desirable.

The original article was received on December 29th, 2012

The revised article was received on July 10th, 2013

The article was accepted on August 28th, 2013

References

- Apple, M. (2000). Can critical pedagogies interrupt rightist policies? *Educational Theory*, 50, 229-254. doi: 10.1111/j.1741-5446.2000.00229.x
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18, 215-228. doi: 10.1080/0268093022000043065
- Biesta, G. J. J. (2004). Against learning. Reclaiming a language for education in an age of learning. *Nordisk Pedagogik*, 24, 70-82.
- Biesta, G. J. J. (2005). The learning democracy? Adult learning and the condition of democratic citizenship. *British Journal of Sociology of Education*, 26, 693-709. doi: 10.1080/01425690500293751
- Biesta, G. J. J. (2006). *Beyond learning: Democratic education for a human future*. Boulder, Co: Paradigm Publishers.
- Biesta, G. J. J. (2009a). Good education in an age of measurement. *Educational Assessment, Evaluation and Accountability*, 21, 33-46. doi: 10.1007/s11092-008-9064-9
- Biesta, G. J. J. (2009b). What kind of citizenship for European Higher Education? Beyond the competent active citizen. *European Educational Research Journal*, 8, 146-157. doi: 10.2304/eeerj.2009.8.2.146
- Biesta, G. J. J. (2010). *Good education in an age of measurement: Ethics, politics, democracy*. Boulder, Co: Paradigm.
- Biesta, G. J. J. (2011). How useful should the university be? On the rise of the global university and the crisis in higher education. *Qui Parle: Critical Humanities and Social Sciences*, 20(1), 35-47. doi: 10.1353/qui.2011.0040
- Biesta, G. J. J. (2013). Interrupting the politics of learning. *Power and Education* 5(1), 4-15. doi: 10.2304/power.2013.5.1.4
- Biesta, G. J. J. (2014). *The beautiful risk of education*. Boulder, CO: Paradigm Publishers.
- Bogotch, I., Mirón, L., & Biesta, G. (2007). Effective for what; Effective for whom? Two questions SESI should not ignore. In T. Townsend (Ed.), *International handbook of school effectiveness and school improvement* (pp. 93-110). Dordrecht/Boston: Springer.
- Charlton, B. G. (1999). The ideology of 'accountability'. *Journal of the Royal College of Physicians of London*, 33, 33-35.
- Charlton, B. G. (2002). Audit, accountability, quality and all that: The growth of managerial technologies in UK Universities. In S. Prickett, & P. Erskine-Hill (Eds.), *Education! Education! Education! - Managerial ethics and the law of unintended consequences* (pp. 13-28). Exeter: Imprint Academic.
- Davis, A., & White, J. (2001). Accountability and school inspection: In defence of audited self-review. *Journal of Philosophy of Education*, 35, 667-681. doi: 10.1111/1467-9752.00252
- Epstein, D. (1993). Defining accountability in education. *British Educational Research Journal*, 19, 243-257.
- Gewirtz, S. (2002). *The managerial school. Post-welfarism and social justice in education*. London/New York: Routledge.
- Granger, D. (2008). No Child Left Behind and the spectacle of failing schools: The mythology of contemporary school reform. *Educational Studies*, 43, 206-228. doi: 10.1080/00131940802117654
- Green, A. (1990). *Education and state formation*. London: Macmillan.
- Henry, G. T. (2002). Choosing criteria to judge program success: A values inquiry. *Evaluation*, 8, 182-204.
- Hess, F. M. (2006). Accountability without angst? Public opinion and No Child Left Behind. *Harvard Educational Review*, 76, 587-610.
- Hopmann, S. (2008). No child, no school, no state left behind: Schooling in the age of accountability. *Journal of Curriculum Studies*, 40, 417-456. doi: 10.1080/00220270801989818
- House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: SAGE.
- Kumashiro, K. K. (2012). *Bad teacher! How blaming teachers distorts the bigger picture*. New York: Teachers College Press.
- Marquand, D. (2004). *Decline of the public*. Cambridge: Polity Press.
- Martens, K., Nagel, A.-K., Windzio, M., & Weymann, A. (Eds.). (2010). *The transformation of education policy*. New York: Palgrave/Macmillan.
- Michie, J. (2011). *The handbook of globalisation. Second edition*. Cheltenham: Edward Elgar.
- Nicolaidou, M., & Ainscow, M. (2005). Understanding failing schools: Perspectives from the inside. *School Effectiveness and School Improvement*, 16, 229-248. doi: 10.1080/09243450500113647

- O'Neill, O. (2002). *BBC Reith lectures 2002. A question of trust*. Retrieved from <http://www.bbc.co.uk/radio4/reith2002>
- Phillips, D. (2005). Policy borrowing in education: Frameworks for analysis. In J. Zajda et al. (Eds.), *International handbook on globalisation, education and policy research: Part 1* (pp. 23-34). Dordrecht: Springer.
- Phillips, D., & Ochs, K. (Eds.). (2004). *Educational policy borrowing: Historical perspectives*. Didcot: Symposium.
- Poulson, L. (1996). Accountability: a key-word in the discourse of educational reform. *Journal of Education Policy*, 11, 579-592.
- Poulson, L. (1998). Accountability, teacher professionalism and education reform in England. *Teacher Development*, 2, 419-432.
- Power, M. (1994). *The audit explosion*. London: Demos.
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford: Oxford University Press.
- Rizvi, F., & Lingard, B. (2009). *Globalizing education policy*. London/New York: Routledge.
- Schwandt, T., & Dahler-Larsen, P. (2006). When evaluation meets the 'rough' ground' in communities. *Evaluation*, 12, 496-505. doi: 10.1177/1356389006073745
- Tomlinson, S. (1997). Sociological perspectives on failing schools. *International Studies in Sociology of Education*, 7, 81-98.
- Townsend, T. (Ed.). (2007). *International handbook of school effectiveness and school improvement*. Dordrecht/Boston: Springer.
- Tröhler, D. (2010). Harmonizing the educational globe. World polity, cultural features, and the challenges to educational research. *Studies in Philosophy and Education* 29(1), 5-17. doi: 10.1007/s11217-009-9155-1
- Usher, R. (2006). Lyotard's performance. *Studies in Philosophy and Education*, 25, 279-288. doi: 10.1007/s11217-006-9009-z
- Waldow, F. (2009). What PISA did and did not do: Germany after the 'PISA-shock' (Review Essay). *European Educational Research Journal*, 8(3), 476-483. doi: 10.2304/eej.2009.8.3.476
- Webb, P. T. (2011). The evolution of accountability. *Journal of Education Policy*, 26(6), 735-356. doi: 10.1080/02680939.2011.587539