

Dimitra Rapti

Postgraduate Student: Comparative Education MA, University College London,
Institute of Education

Module: Comparative Education: Theories and Methods

To what extent should education policy be influenced by international comparisons of pupil achievement?**Introduction**

Over the past few decades, educational assessment has witnessed remarkable progress through the introduction of new ways to assess pupil achievement and it has been largely viewed as an aspect of policy consideration (Maclean and Watanabe, 2010). In that context, since the 1960s, international organizations (IOs) have conducted international comparative studies of student performance, which have raised the concept of benchmarking in education and have influenced significantly the global education policy (Berry and Adamson, 2011, p.8).

The adherence of the field of Comparative Education to politics dates back to the nineteenth century, when the use of scientific evidence by comparative studies became the basis of educational reforms (Cowen, 2006, p.563). Nowadays, this idea for many still constitutes the driving force of Comparative Education (ibid.). The difference between the previous pursuits and today's aspirations of the field is the emergent importance of considering student achievement as an explanatory tool of the educational and economic reality (Auld and Morris, 2013, p.130).

In this critical review I will evaluate the degree to which international comparisons of pupil achievement affect education policy, the implied factors that shape them, as well as their impact on forming "indicators of quality" in education; indicators that prescribe the actions of policy makers. The connection of international assessments to education policy is a controversial issue that poses a series of questions, which constitute the reference points of this paper. To what degree do they provide robust evidence, which could be infused in the process of policy-making? Should they contribute to the formation of education policy and reform or would they constitute a threat to the multidimensional role of education? In the final analysis, to what extent is the reinforcement of education quality ensured?

These are questions that have incited a variety of arguments and provoked a debate over the use of international tests in education policy. Some stand in the methodological side of the argumentation, by analysing the strengths and limitations of the international assessments' outcomes, while others reflect upon the implications of their use in education policy.

By taking into account these arguments, I will assess the extent to which comparative studies of student performance should influence education policy and I will conclude that although they should have an informative role on educational decisions, they should not constitute the baseline of a nation's educational policy direction. These arguments will be explored in greater depth in the following sections, after discussing the growth of international comparisons in the context of globalization.

1. Globalization & The Growth of International Assessments

Education had always been seen as the sector most adherent to national activity (Dale and Robertson, 2007, p.2), but as Kamens (2013) notes, national influence on education policy has recently been diminished by the power of global goals. Nations are incrementally conceptualized as “open systems” towards the growth of globalization (ibid.).

The concept of “global governance in education” derives from the neoliberal ideas developed in the 1980s and the subsequent turn to “New Public Management” (Auld, 2014). In this context, research on the reinforcement of “educational effectiveness” -a forerunner of the focus on measuring educational outcomes- through “evidence-based” knowledge brought about a long educational discourse (Biesta, 2008, p.34). Evaluations fuelled the educational sector and were used as a means of establishing educational changes (Lundgren, 2011, p.18). The turn towards a “global knowledge economy” and the need for high quality educational data has transferred these practices to the international scene (Auld, 2014).

Ushering in a new globalized era, IOs have gained the capacity of driving nations politically and socially (Finnemore, 1996; Barnett & Finnemore, 1999, cited in Meyer and Benavot, 2013, p.12). Over the recent years, they have penetrated every aspect of the educational sector on a global level, which has led to an incrementing “internationalization” of education (Pereyra, Kotthoff and Cowen, 2011, p.2). International organizations’ most dynamic and powerful impact on education is their engagement with the development of comparative indicators for the purpose of evaluating the various educational systems (Dale and Robertson, 2007, p.6). In this respect, recently, there has been an exponential increase of the use of international assessments that measure student performance across nations, led by IOs (Kirsch et al., 2013, p.1).

International large-scale assessments, in their broad term, are surveys that measure comparatively the skills of a target group in a specific area (ibid.). They are based on the fact that skills are considered to be a prerequisite for development and concentrate on a population’s overall outcomes, rather than individual results (ibid.). International agencies, such as the International Association for the Evaluation of Educational Achievement (IEA) and the Organization for Economic Cooperation and Development (OECD), undertake such comparative studies in educational achievement and their outcomes influence worldwide educational activities and policy plans (Crossley, 2002, cited in Nóvoa and Yariv-Mashal, 2003, p.425).

In detail, the IEA’s Trends in International Mathematics and Science Study (TIMSS) measures trends in mathematical and scientific performance of the fourth and eighth grade students every four years (TIMSS & PIRLS, International Study Center, n.d.). The Progress in International Reading Literacy Study (PIRLS) is also a study run by IEA, which measures trends of the fourth grade in reading performance every five years (ibid.). Furthermore, the IEA has conducted some studies related to civic education, such as the Civic Education Study (CIVED) of 1999, which investigated the civic knowledge and perceptions of secondary and upper secondary school pupils, and the International Civic and Citizenship Education Study (ICCS) of 2009 that examined the extent to which educational systems prepare young people for their role as citizens in the twentieth first century (Schulz et al., 2010).

Finally, the OECD’s Programme for International Student Assessment

(PISA) assesses fifteen-year-old pupils in reading, science and mathematics (National Center for Education Statistics, n.d.). It is conducted on a three-year rotation and in each round of testing it gives prominence to one of the aforementioned domains (ibid.). The PISA project does not test pupils' curriculum attainment, but draws attention to the application of their knowledge in different contexts of everyday life (Murat and Rocher, 2004, p.190). On the contrary, TIMSS intends to investigate the students' educational outcomes with regards to the curricula (ibid.).

2. International Comparisons & Education Policy

Today there is a shift of the global education policy to the significance of educational outcomes and, by extension, to the effectiveness of educational systems (Hanushek and Woessman, 2010, cited in Valverde, 2014, p.576). In this regard, comparisons of the outcomes of international assessments stand in the forefront of current educational activities, as there is an increasing demand by policy makers to draw evidence that could be used as a means of enhancing the educational sector (Gil, 1999, cited in Owen et al., 2004).

However, these studies do not play merely an informative role in education policy. The inferences drawn by IOs through their comparative studies are central to the global educational policy discourse and constitute the baseline of governments' policy plans (Crossley, 2002, cited in Nóvoa and Yariv-Mashal, 2003, p.425). This political engagement of the comparative studies in education results in the development of policies by specific "standards, outcomes and benchmarks" (ibid., p.427).

In detail, such surveys generate hierarchies of effective and ineffective educational systems and use them as a basis to draw conclusions and recommend ways for the improvement of educational quality (ibid., p.425). These hierarchies are presented through league tables that place countries on a scale of performance, regarding the scores of their students on the assessment (Murat and Rocher, 2004, p.190). As Grek (2013, p.707) notes, the universally accepted knowledge -derived from the assessments' results- that experts promote seems to be turned directly into policy. According to Nóvoa (2002, p. 144, cited in Grek, 2010), "comparing must not be seen as a method, but as a policy ... the expert discourse builds its proposals through 'comparative' strategies that tend to impose 'naturally' similar answers in the different national settings".

The high degree to which these studies affect education policy is demonstrated by the fact that there are cases where low rankings have brought about a "shock" in the educational reality of some nations (Morris, 2011, p.4). The release of the 2000 PISA results in Germany, for instance, has provoked a lot of tension in the German society (Wiseman, 2013, p.304). In response to such cases, policy actors turn to the adoption of policies that are identified to be successful in other countries (Morris, 2011, p.4). The impact of these studies is magnified and becomes more appealing by the dissemination of their results through the media (Nóvoa and Yariv-Mashal, 2003, p.425). This makes it even more imperative for actions to be taken by governments instantly to "cover the gaps" of their educational system by aligning their decisions with unquestionable, widely embraced practices (ibid.).

Building on the above ideas, it could be argued that over the past few years a new trend of "education governance" has been making its first steps, driven by IOs

(Grek, 2010). As Jones (2007, p.94, cited in Grek, 2010) points out, IOs have functioned as “purveyors of ideas” and Grek (2010, p.396) adds that they have reduced significantly the national autonomy on education policy.

This idea has been extended by Valverde (2014, p.576), who claims that standards of quality are defined by IOs. He suggests the term “radical advocates of education”, as well as “international regimes”, for all the organizations that drive nations to adjust their education policy to “internationally derived definitions of educational quality”. This tendency of IOs has been further explained by Dale and Robertson (2007), who maintain that the fact that education is a broad sector and does not entail any specific definition allows IOs to develop their own indicators of quality.

OECD has gained ground in the field of education and stands at a dominant position among other IOs, because of its principal driving force, the so-called “comparative turn” - “a scientific approach to political decision making” (Martens 2007, cited in Grek, 2009, p.25). PISA is one of the means by which OECD is making its greatest contribution in the area of education policy (Grek, 2009, p.25) and it is regarded to be “the most successful enterprise in Comparative Education” (Hopmann and Brinek, 2007). According to OECD (2004, p.12, cited in Wiseman, 2013), PISA’s main role is to reflect on the policy-making process of the educational sector by satisfying the need of countries to gain “policy lessons”.

In an attempt to evaluate the impact of PISA, an example could be drawn by Lawn and Grek (2012, p. 133), who investigated its placement in Europe and argued that its dataset is used as an explanatory tool of any reforms or ongoing policies. In some cases, changes in education that otherwise would be disregarded, can be established under the flag of PISA (*ibid.*, p.134).

3. International Assessments as a Useful Tool in Education Policy

International tests of student performance could be regarded as a valuable source of information in several ways. According to Klieme (2013, p.116), international assessments offer indicators of evaluating educational systems and play a vital role in the understanding of “educational effectiveness”. He mentions that these studies are linked to policy-making by providing information on how educational goals were selected and prioritized in the various educational systems, identifying elements that could be controlled by a system’s administration and by offering “best practices” of specific systems that could be seen as reference points by other countries.

Lockheed (2013) supports the opinion that cross-national studies could draw attention to some areas of education policy that could not be evaluated with a similar effectiveness nationally. For instance, she suggests that curricular reforms, as well as attempts to ameliorate teaching methods, could be more easily stimulated by comparing a country’s practices with those of countries that achieved higher scores. In addition to that, Lockheed argues that investment practices are also a domain of the education policy in which international studies could be productive. Reporting the results of countries with equal economic standards, but with different scores in the assessments, could function as an incentive for the low-performing countries to amend their investment practices (*ibid.*).

Moreover, Lockheed (2013) concentrates on the impact of international assessments in developing countries and concludes that to some degree they have

influenced them in a positive way. Participating in international testing has been proved beneficial not only in reinforcing their educational system by affecting the areas of curriculum and teaching practice, but also by enhancing their capacity for assessment (ibid.).

Another thing worth mentioning is the fact that international assessments provide a valuable source of information to countries in which there is no national examination or to countries with a decentralized educational system that do not have a prescribed national curriculum. These nations could be benefited through international tests, as they could obtain a sense of their students' educational outcomes (Owen et al., 2004, p.10-11).

Drawing on the example of PISA, Owens (2013, p.43) proposes that on condition that researchers deal with the test's outcomes rigorously, they could contribute greatly to the advancement of social science fields. In parallel, she stresses the great amount of information that could be drawn from PISA data beyond student performance -inequalities, gender issues, etc.- that could enhance the world's perception of education internationally (p. 27 and 43).

In this context, PISA dataset can be used as a resource to highlight inequalities within a system that would have not been noticed otherwise. For instance, England -a country in which student performance has been found to be largely related to students' socio-economic status- has been further explored by Jerrim (2012) using data from PISA in order to reveal how achievement is distributed among the student population within the country. Similar findings could serve as an advantage to education policy by revealing aspects of education that require a supplementary attention by the side of governments.

4. Limitations of International Assessments' Data in Education Policy

According to Lawn and Segerholm (2011, p.45), "quantitative data produce new global norms about standards and measuring". Taking into account the significance of data in today's world, a question that arises is how reliable the evidence that international surveys provide it really is. Next, several arguments and studies are presented, which reveal a variety of limitations emerged from international tests that question the data itself, as well as its interpretation. These studies consider the weaknesses of international comparisons on forming league tables that direct world-wide educational policy decisions.

i. Reliability of Data and Rankings

Considering the case of PISA, Hopmann and Brinek (2007) stress the fact that the study is characterized by cultural biases and methodological limitations, which do not allow a clear picture of each country's educational reality and schooling to be framed.

Regarding the TIMSS test, variations of the countries' curricula reduce the assessment's objectivity and accuracy, even though it is designed to examine only the common items of the curricula (Ashcraft, 2009). The limitations of the international assessments become even more apparent when they generate different outcomes in the same country. The United States, for example, came twenty-ninth in the rankings of the 2009 PISA test in numeracy, while it reached the ninth place of the 2011 TIMSS test in mathematics (Buckingham, 2012, p.3).

Rankings are also criticized in terms of their reliability to guide education policy. As it is stressed by Tomlinson (2014, cited in Wilby, 2014a), the size and the context of each country appeared on the PISA league tables are not easily comparable. Another

noteworthy aspect is that forming rankings -which “judge” the effectiveness of educational systems- based solely on the tests’ outcomes could provide unreliable evidence, as student performance is not affected only by a country’s school environment. Meyer and Schiller (2013, p.207) claim that external factors to the educational process in schools, particularly social, economic and cultural, influence substantially PISA results. This argument raises doubts concerning conclusions drawn by rankings that seem to “blame” or “praise” the success of systems to provide education. This is because the educational system alone could not account for the overall educational situation in a country.

An illustrative example is the case of Finland; a country that has always stood on the top of the PISA scores (Buckingham, 2012, p.3). It is argued that out-of-school aspects may influence its high performance on PISA, indicatively its high levels of social equity, its minimal number of pupils to whom Finnish is not their first language and the fact that the Finnish language is less complex than other languages, which makes its acquisition from students less problematic (*ibid.*, p.1). Considering this example, it could be inferred that rankings developed by tests’ results and used as indicators of an educational system’s success could lack of reliability.

ii. Interpretation of Rankings

Data of student performance is hugely embedded into the agenda of educational administrators, but a cautious translation of the evidence is often left behind (Earl and Louis, 2013, p.200-203). Considering the fact that the quality of a decision is based on the quality of the available data, a potential misuse in the interpretation of the collected evidence could lead to the formation of a misleading global view (*ibid.*).

The 2007 and 2010 reports by the “McKinsey and Company”¹, which, as Coffield (2012) mentions, have strongly affected education policy, could serve as an example. The reports translate data from international and national surveys into information that can be used by policy makers to gain an insight into successful practices from high-performing countries and improve their systems (Mourshed, Chijioke and Barber, 2010). However, by examining the first report, Alexander (2010) highlights the ambiguous way that it is presented. Among Coffield’s (2012) critical arguments against their reliability, stands their superficiality, the underestimation of the cultural factor and the technocratic way of their approach.

A number of studies have challenged the inferences drawn by rankings regarding the identification of “world class” schools. Zhao and Meyer (2013), for instance, investigated if the promotion of the “creative thinking” - a significant aspect of economic growth - through the educational system is an element of the PISA high-achieving East Asian countries. However, they found out that this feature is suppressed by their educational culture. On the contrary, discipline and submissiveness are the main cultural aspects that drive their system. This implicates that high ranking in standardized tests does not ensure success in every aspect that education could aim to.

A similar view is held by Heyneman (2013), who suggests that East Asian

¹ “How the world’s best-performing school systems come out on top” and “How the world’s most improved school systems keep getting better” are the titles of the 2007 and 2010 McKinsey’s reports respectively (Mourshed, Chijioke and Barber, 2010).

countries should not be seen as a role model in education, as students of these countries suffer from stress and endless out-of-school tutoring in order to succeed in school. Rankings fail to disclose those aspects of the educational reality of high-performing countries. Consequently, education policy guided by conclusions of international studies is subject to misleading interpretations of a “successful” educational system’s characteristics.

All in all, Biesta (2008, p.35) points out that policy produced by conclusions drawn by league-table comparisons is subject to any limitations that these surveys entail; unreliable data or misleading inferences. He explains that the anticipated limitations do not only refer to the lack of “technical validity”, but also to the lack of “normative validity” of the measurements. In other words, to the possibility of getting distracted by unreliable data is added the risk of getting carried away by overestimating the indicators of performance and placing them as a principal goal. In this way, the initial purpose is neglected and therefore, “indicators of quality become mistaken for quality itself” (ibid.).

5. Implications in Education Policy

As detailed below, there is a large amount of literature produced to address the consequences of using the outcomes of international studies in the formation of education policy. Some arguments reflect upon the implied intentions regarding their use by both IOs and governments. Other arguments allude to the anticipated direction that education will be driven to follow by concentrating on the “standards of quality” that these studies promote.

i. The Nature of the “Standards of Quality”

As it has been explained previously, education policy is greatly affected by the “standards of quality” that IOs have founded through their studies. Hence, in order to discuss their impact on education policy, it would be necessary to take a closer look at the implicated factors that surround the purposes of IOs through the educational assessments and, by extension, the nature of the standards themselves.

According to Nóvoa and Yariv - Mashal (2003, p.424), the need to develop comparative indicators that enable the measurement of the “efficiency” and “quality” of education is grounded in the increasing interconnection of economic competition with educational goals. In addition to that, as Dale and Robertson (2007) argue, IOs are driven by the idea of the significance of education as a principal determinant of economic prosperity.

In that perspective, Grek (2009, p.27) holds the view that PISA is concerned about the assessment of skills in domains that are prioritized in a market-oriented society: reading, mathematics and science. Other “secondary” skills, such as the arts and citizenship, are undermined (ibid.). Considering the fact that OECD is an economic and market-oriented organization per se (Meyer and Benavot, 2013), it could be argued that the nature of the skills assessed by PISA corresponds to the nature of OECD’s orientation in education policy; an orientation shaped by economic factors. Consequently, the “indicators of quality” established by PISA refer to a narrow range of skills, which are underpinned by a technocratic view of education.

Thus, it could be inferred that education policy shaped under the indicators

developed by IOs is mainly directed by economic intentions. This direction increases the importance of a skill-based education -a prerequisite for the advancement of economy (Alexander, 2010, p.804)- which could function as a constraint towards the overall role of education in society.

ii. Focus on Standards & Rankings

As it has been noted previously, governments concentrate their education policy on rankings and standards set by IOs. This focus on rankings raises questions of the extent to which policy makers eventually proceed to reforms that ‘matter’ for the enhancement of their educational system, as well as the degree to which meeting the prescribed educational standards that rankings suggest, would reinforce quality in education.

On one hand, a reason why governments give prominence to the rankings is highlighted by Morris (2011). As he claims, to some extent, the use of comparisons with high-achieving educational systems as undoubtable information of what is considered to be successful allows governments to promote the policies that suit their agendas. As it is indicated by the 2010 White Paper of the English government and its main points of reference, such as the McKinsey reports, there is a tendency of selectivity in the use of evidence, as well as “mixing and matching” information, which lacks of objectivity (ibid.). Consequently, in these cases, reforms made upon selective and subjective information could not provide a solid ground for educational advancement to take place.

On the other hand, a reason why governments use the rankings as a reference point of their policies is to reach a higher place in the league tables and outperform the rest of the countries (Alexander, 2010; Biesta, 2008). In detail, Alexander (2010, p. 801) states that an increasing number of wealthy countries, having already solved the educational problems that developing countries struggle to address, take part in a competition pursuing a “world class” position in the educational scene. The reason why governments have raised a strong interest in comparing the rankings of their educational systems with those of their competitors is justified by Valverde’s (2014, p.576) stance that policy makers relate the productivity of their educational systems with the market-based global competition. As Valverde (2014) explains, investing in educational effectiveness is an imperative for a “knowledge economy”.

According to Biesta (2008, p.34), international comparisons of student achievement foster this competition, as they are portrayed through league tables. The so-called “league-table studies” aim to raise nations’ awareness on their position in the global scale and their inferences are used as “evidence-based” information that guides the planning of education policy in the name of a “raising standards” need (ibid.). In this context, educational experts develop sets of actions that could be adopted by policy makers in order to lead their country to a higher place in the rankings and therefore ensure their nation’s economic prosperity (Auld and Morris, 2013).

An indicative example that illustrates this need for “winning” is the response of the English government after the diffusion of the PISA outcomes in 2010: “The truth is, at the moment we are standing still while others race past... We have no choice but to be this radical if our ambition is to be world-class” (Department for Education, 2010, p.3-4). In this regard, Alexander’s (2012) words illustrate the

competitive dimension imposed to today's global education: "Like all races, I need hardly add, the global education race allows only two outcomes, winning or losing".

In response to the focus of governments to follow this "race", some arguments have been raised concerning the degree to which they eventually benefit their educational system. As it is stressed by Wilby (2014a), for instance, education policy actors anticipate anxiously the outcomes of PISA with the hope that the students of their nation have outperformed the rest of the world. With the spotlight placed on the rankings, approximately half of the participated countries have proceeded to considerable educational reforms (ibid.). However, they often resort to changes that do not have the potential to be proved beneficial in the long run, but they function as a quick solution, just in the hope of achieving a higher place in the next round of rankings (Wilby, 2014b).

All in all, as Ball (2014) maintains, the education world is nowadays under the control of numbers and rankings, which determine the conceptualization of education policies. In this respect, Valverde (2014, p.585) points out the lack of "empirical verification" that achieving the proposing educational standards – inferred by the rankings – reinforces the quality of education. In the same vein, as Alexander (2010, p.801) mentions, the fact that "world class" is mainly shaped by the outcomes of the international assessments implies that reaching this worldwide recognition, entails the risk of creating educational systems that emphasize narrowly the notion of learning and undervalue the educational "quality, equity and governance"; aspects that define a system which is "civilised as well as competitive".

iii. Consequences on Education

For Valverde (2014, p.576) quality today is connected to educational outcomes and as Ozga et al. (2011, p.2) note, "quality is conformity with standards". However, as it is indicated below, the proposed "standards of quality" of international assessments have drawn attention for criticism about the way "quality" is perceived. Moreover, central to the educational discussion have been the consequences of following these indicators on the role that education is supposed to play.

Olseen (2004, cited in Meyer and Benavot, 2013, p.12) argues that in the process of reaching the "indicators of quality", education is shifted from its concentration on national cultural and social aims to the pursuit of preparing students for the workforce. This shift that links education with economic intentions constitutes a threat to the civic orientation of education and, by extension, to the promotion of democratic values (ibid.). A similar stance about the loss of the democratic perspective of education is taken by Grubb and Lazerson (2006, p.301, cited in Meyer and Benavot, 2013, p.12). As Slee and Weiner (2000, cited in Townsend, 2001, p.121) claim, school effectiveness is simply concentrated on the "mechanics of schooling" and neglects to denote the "aims of schooling".

Moreover, teaching is shaped by a variety of factors that transform it into a complex procedure (Alexander, 2008, p.19). In essence, pedagogy is more than a technical issue and thus it cannot be treated as a series of "inputs", "processes" and "outcomes" (ibid., p.20). As it was stated by the founders of IEA, "treating the world as a laboratory" serves as a way to identify educational problems and successful practices as solutions (Murat and Rocher, 2004, p.190). Accordingly, OECD aims to investigate deeply the educational effectiveness by measuring the "resources (input)

and pupil outcomes (output)” (ibid.).

However, the complex pedagogical issues that cannot be monitored by indicators should also be taken into account (Alexander, 2008, p.43). As Alexander (2010, p.813) highlights, there is a tendency to “reduce quality to quantity in order that it can be indicated and measured”. Any aspect of education that deviates from the prescribed educational objectives is seen as an abnormality that makes education “difficult”, but this “difficulty” is what defines education after all (Biesta, 2001, p.386).

In that respect, a large amount of academics from twelve countries have recently sent a letter to OECD to express their concern on the extent to which league tables of PISA influence education negatively, by giving prominence to the fact that reforms should not be relied on “a single narrow measure of quality” (The Guardian, 2014). They claim that PISA “impoverishes our classrooms”, by highlighting the alienation of education from the multidimensional development and wellbeing of children that it should have aimed to (ibid.).

Finally, the importance of the proposed standards on shaping our picture of education quality is also criticized. Biesta (2008) emphasizes that although the growth of the “measurement culture” in education to some degree is fruitful, it insinuates that all educational activities should be determined by the produced “evidence-based” information. He expresses his concerns that the purpose of education is all too often disregarded and highlights the need to couple educational decision-making with a “value-based” evaluation of the decisions (ibid.). In that context, he proposes a turn from an “evidence-based education” to a “value-based education” (Biesta, 2010).

Conclusion

All things considered, it is not the international tests’ data itself that constitutes a threat to the future of education, but the way it is often promoted by IOs and used by governments. To my point of view, international studies of student achievement should only be seen as a supplementary tool by policy actors in education. In this way, they could play a vital role towards the advancement of some aspects of the educational systems. After all, educational outcomes based on numbers could never be regarded as a panacea that could lead any educational decision.

In other words, education policy could benefit from the knowledge that these surveys offer, provided that it is used cautiously. This means that policy makers should detach their educational decisions from the goal of achieving a higher place in the rankings that comparative studies promote. It is also very important that the limitations of these studies, regarding their methodology, their measurements and their impact on a potential reform that they could suggest, are to be taken into account during the policy-making process.

Looking more closely at the positive side of the spectrum, with regards to the rankings produced by international standardized tests, I would suggest that even if they can be misleading to some degree, they could be considered beneficial in the process of identifying educational inequalities or dysfunctional policy directions, to which no attention has been previously paid on a national level. This could prove a valuable application of the enormous datasets that these tests offer, as they can provide a useful catalyst for changes in cases where education is intentionally or unintentionally suppressed. In this context, educational outcomes of large-scale assessments could build momentum for beneficial policy reform.

On the other hand, overvaluing the significance of the rankings and translating them into levels of educational quality, by ignoring the fact that they mirror a narrow and often an unreliable dimension of the educational reality, could prove disastrous for the role that education will play in the future. Aligning education policy with objectives that are solely derived from the international benchmarking, automatically entails an acceptance of the fact that education should aim to economic and market-oriented goals; objectives that the “indicators of quality” of these studies promote. Biesta’s words feature precisely the significance of understanding what we value in education: “Whether a high score on TIMSS, PIRLS or PISA does indeed indicate good education is an entirely open question that crucially depends on what we expect from education...” (Biesta, 2009).

Finally, as Alexander (2010, p.816) suggests, an “educational consciousness”, which does not only correspond to the economic pursuits of education, but it is also defined morally, is a requirement in today’s globalized world. This is the only way to achieve true “world class education” (ibid.).

References

- Alexander, R. (2008). *Education for all, the quality imperative and the problem of pedagogy. Research Monograph No 20*. CREATE, Institute of Education, University of London.
- Alexander, R. (2010). ‘World class schools’ noble aspiration or globalised hokum?. *Compare: A Journal of Comparative and International Education*, [online] 40(6), pp. 801-817.
Available at:
<http://www.tandfonline.com/doi/abs/10.1080/03057925.2010.523252#.VLwZDEesURo> [Accessed 15 Dec. 2014].
- Alexander, R. (2012). *Visions of Education, Roads to Reform: The global educational race and the Cambridge Primary Review*. (Lecture). Universidad Gabriela Mistral, Santiago, Chile.
Available at: http://cprtrust.org/wp-content/uploads/2014/06/Alexander_Chile_GMU_lecture_1_121030.pdf [Accessed 16 Dec. 2014].
- Ashcraft, C. (2009). *Comparing U.S. K-12 Students’ Math and Science Performance Internationally. What are the facts, what do they mean for educational reform, and how do I talk effectively about the issues?*. [online] National Center for Women & Information Technology (NCWIT). Available at: <http://www.ncwit.org/sites/default/files/resources/talkingpoints-mathscience.pdf> [Accessed 17 Feb. 2015].
- Auld, E. (2014). Development Inc. and the Ghost in the Machine: The Idea of Global Governance. *Norrag News*, [online] (51), p. 27.
Available at:
<http://www.norrag.org/fileadmin/Full%20Versions/NN51.pdf> [Accessed 11 Jan. 2015].
- Auld, E. and Morris, P. (2013). Comparative education, the ‘New Paradigm’ and policy borrowing: constructing knowledge for educational reform. *Comparative Education*, [online] 50(2), pp. 129-155.
Available at:

- <http://www.tandfonline.com/doi/abs/10.1080/03050068.2013.826497#.VLvizke sURo> [Accessed 14 Dec. 2014].
- Ball, S. (2014). Education and the tyranny of numbers. [Blog] *IOE London Blog*. Available at: <https://ioelondonblog.wordpress.com/2014/11/16/education-and-the-tyranny-of-numbers/> [Accessed 13 Jan. 2015].
- Berry, R. and Adamson, B. (2011). Assessment Reform, Past, Present and Future. In: R. Berry and B. Adamson, eds. (2011). *Assessment reform in education*. Dordrecht: Springer.
- Biesta, G. (2001). How difficult should education be?. *Educational Theory*, [online] 51(4), pp. 385-400. Available at: <http://onlinelibrary.wiley.com/doi/10.1111/j.1741-5446.2001.00385.x/abstract> [Accessed 15 Jan. 2015].
- Biesta, G. (2008). Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, [online] 21(1), pp. 33-46. Available at: <http://link.springer.com/article/10.1007%2Fs11092-008-9064-9> [Accessed 22 Dec. 2014].
- Biesta, G. (2009). *Good Education: What it is and why we need it*. (Lecture). The Stirling Institute of Education. Available at: <https://www.scribd.com/doc/242597554/Biesta-2009-Good-Education-What-It-is-and-Why-We-Need-It> [Accessed 12 Jan. 2015].
- Biesta, G. (2010). Why 'What Works' Still Won't Work: From Evidence-Based Education to Value-Based Education. *Studies in Philosophy and Education*, [online] 29(5), pp. 491-503. Available at: <http://link.springer.com/article/10.1007%2Fs11217-010-9191-x> [Accessed 22 Dec. 2014].
- Buckingham, J. (2012). *Keeping PISA in Perspective: Why Australian Education Policy Should Not Be Driven by International Test Results*. Issue Analysis No. 136. [online] The Centre for Independent Studies (Australia). Available at: <http://www.cis.org.au/images/stories/issue-analysis/ia136.pdf> [Accessed 19 Feb. 2015].
- Coffield, F. (2012). Why the McKinsey reports will not improve school systems. *Journal of Education Policy*, [online] 27(1), pp.131-149. Available at: <http://www.tandfonline.com/doi/abs/10.1080/02680939.2011.623243#.VLwaKE esURo> [Accessed 15 Dec. 2014].
- Cowen, R. (2006). Acting comparatively upon the educational world: puzzles and possibilities. *Oxford Review of Education*, [online] 32(5), pp.561-573. Available at: <http://www.jstor.org/stable/4618682> [Accessed 15 Dec. 2014].
- Dale, R. and Robertson, S.L. (2007). New arenas of global governance and international organisations: reflections and directions. In: K. Martens, A. Rusconi and K. Lutz, eds. (2007). *Transformations of the State and Global Governance*. London: Routledge.
- Department for Education, (2010). *The Importance of Teaching: The Schools White Paper 2010*. London: Crown.
- Earl, L. and Louis, K. S. (2013). Data Use: Where to from here?. In: K. Schildkamp, M. K. Lai and L. Earl, eds. (2013). *Data-based decision making in education*.

- Dordrecht: Springer.
- Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, [online] 24(1), pp. 23-37.
Available at:
<http://www.tandfonline.com/doi/abs/10.1080/02680930802412669#.VLvhjEesURo> [Accessed 15 Dec. 2014].
- Grek, S. (2010). International Organisations and the Shared Construction of Policy 'Problems': problematisation and change in education governance in Europe. *European Educational Research Journal*, [online] 9(3), pp. 96-406.
Available at:
<http://www.worlds.co.uk/rss/abstract.asp?j=eerj&aid=4122&doi=1> [Accessed 16 Dec. 2014].
- Grek, S. (2013). Expert moves: international comparative testing and the rise of expertocracy. *Journal of Education Policy*, [online] 28(5), pp.695-709. Available at:
<http://www.tandfonline.com/doi/abs/10.1080/02680939.2012.758825#.VLw3DUesURo> [Accessed 15 Dec. 2014].
- Heyneman, S. (2013). The International Efficiency of American Education: the bad and the not-so-bad news. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.
- Hopmann, S. and Brinek, G. (2007). Introduction: PISA According to PISA – Does PISA keep what it promises?. In: S. Hopmann, G. Brinek and M. Retzl, eds. (2007). *PISA zufolge PISA - PISA According to PISA*. Wien: Lit.
- Jerrim, J. (2012). The Socio-Economic Gradient in Teenagers' Reading Skills: How Does England Compare with Other Countries?. *Fiscal Studies*, 33(2), pp.159-184.
- Kamens, D. (2013). Globalization and the Emergence of an Audit Culture: PISA and the search for 'best practices' and magic bullets. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.
- Kirsch, I., Lennon, M., Davier, M., Gonzalez, E., Yamamoto, K. (2013). On the growing importance of International Large-Scale Assessments. In: M. Davier, E. Gonzalez, I. Kirsch, and K. Yamamoto, eds. (2013). *The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research*. Dordrecht: Springer.
- Klieme, E. (2013). The role of Large-Scale Assessments in Research on Educational Effectiveness and School Development. In: M. Davier, E. Gonzalez, I. Kirsch, and K. Yamamoto, eds. (2013). *The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research*. Dordrecht: Springer.
- Lawn, M. and Grek, S. (2012). *Europeanizing education*. Oxford: Symposium Books.
- Lawn, M. and Segerholm, C. (2011). Europe through experts and technologies. In: J. Ozga, P. Dahler - Larsen, C. Segerholm and H. Simola, eds. (2011). *Fabricating Quality in Education*, 1st ed. Abingdon, Oxon: Routledge.
- Lockheed, M. (2013). Causes and Consequences of International Assessments in Developing Countries. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.

- Lundgren, U. (2011). PISA as a Political instrument: One History Behind the Formulating of the PISA Programme. In: M. Pereyra, H. Kotthoff and R. Cowen, eds. (2011). *PISA Under Examination: Changing Knowledge, Changing Tests, and Changing Schools*. [ebook] Rotterdam: Sense Publishers. Available at: <http://www.cese-europe.org/images/cese/general/pisa%20under%20examination.pdf> [Accessed 14 Dec. 2014].
- Maclean, R. and Watanabe, R. (2010). Foreword. In: R. Berry and B. Adamson, eds. (2011). *Assessment reform in education*. Dordrecht: Springer.
- Meyer, H. D. and Benavot, A. eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.
- Meyer, H. D. and Schiller, K. (2013). Gauging the Role of Non-educational Effects in Large-scale Assessments: socio-economics, culture and PISA outcomes. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.
- Morris, P. (2011). Pick n' mix, select and project; policy borrowing and the quest for 'world class' schooling: an analysis of the 2010 schools White Paper. *Journal of Education Policy*, [online] 27(1), pp.89-107.
Available at:
<http://www.tandfonline.com/doi/abs/10.1080/02680939.2011.596226#.VLvjHEesURo> [Accessed 15 Dec. 2014].
- Mourshed M., Chijioke, C. and Barber, M. (2010). *How the world's most improved school systems keep getting better*. [online]
Available at: <http://mckinseysociety.com/how-the-worlds-most-improved-school-systems-keep-getting-better/> [Accessed 15 Dec. 2014].
- Murat, F. and Rocher, T. (2004). The methods used for international assessments of educational competences. In: J. Moskowitz and M. Stephens, eds. (2004). *Comparing learning outcomes*. London: RoutledgeFalmer.
- National Center for Education Statistics, (n.d.). *Program for International Student Assessment (PISA) - Overview*. [online]
Available at: <http://nces.ed.gov/Surveys/PISA/> [Accessed 14 Jan. 2015].
- Nóvoa, A. and Yariv-Mashal, T. (2003). Comparative Research in Education: a mode of governance or a historical journey?. *Comparative Education*, [online] 39(4), pp.423-438.
Available at:
<http://www.tandfonline.com/doi/abs/10.1080/0305006032000162002#.VLvjSkeSURo> [Accessed 15 Dec. 2014].
- Owen, E., Stephens, M., Moskowitz, J. and Gil, G. (2004). Toward education improvement: The future of international assessment. In: J. Moskowitz and M. Stephens, eds. *Comparing learning outcomes*. London: RoutledgeFalmer.
- Owens, T. (2013). Thinking beyond League Tables: a review of key PISA research questions. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.
- Ozga, J., Dahler-Larsen, P., Segerholm, C. and Simola, H. eds. (2011). *Fabricating Quality in Education*. 1st ed. Abingdon, Oxon: Routledge.
- Pereyra, M., Kotthoff, H. and Cowen, R. eds. (2011). *PISA Under Examination: Changing Knowledge, Changing Tests, and Changing Schools*. [ebook]

- Rotterdam: Sense Publishers.
Available at:
<http://www.ceseeurope.org/images/cese/general/pisa%20under%20examination.pdf> [Accessed 14 Dec. 2014].
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D. and Losito, B. (2010). *Initial Findings from the IEA International Civic and Citizenship Education Study*. [online] Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).
Available at:
http://www.mineducacion.gov.co/1621/articles-246644_archivo_pdf_ICCS.pdf [Accessed 7 Feb. 2015].
- The Guardian, (2014). *OECD and Pisa tests are damaging education worldwide - academics*. [online]
Available at: <http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics> [Accessed 13 Jan. 2015].
- TIMSS & PIRLS, International Study Center, (n.d.). *About TIMSS and PIRLS*. [online] Available at: http://timssandpirls.bc.edu/home/pdf/TP_About.pdf [Accessed 17 Jan. 2015].
- Townsend, T. (2001). Satan or Saviour? An Analysis of Two Decades of School Effectiveness Research. *School Effectiveness and School Improvement*, [online] 12(1), pp.115-129.
Available at:
<http://www.tandfonline.com/doi/abs/10.1076/sesi.12.1.115.3461#.VLxAQkesURo> [Accessed 22 Dec. 2014].
- Valverde, G. (2014). Educational Quality: Global Politics, Comparative Inquiry, and Opportunities to Learn. *Comparative Education Review*, [online] 58(4), pp.575-589. Available at:
<http://www.jstor.org/discover/10.1086/678038?uid=3738032&uid=2&uid=4&sid=21105100406451> [Accessed 22 Dec. 2014].
- Wilby, P. (2014a). *Academics warn international school league tables are killing 'joy of learning'*. [online] The Guardian.
Available at: <http://www.theguardian.com/education/2014/may/06/academics-international-school-league-tables-killing-joy-of-learning> [Accessed 12 Jan. 2015].
- Wilby, P. (2014b). *Global school tests under attack as OECD accused of killing 'joy of learning'*. [online] The Guardian.
Available at: <http://www.theguardian.com/business/2014/may/06/oecd-pisa-tests-education-joy-of-learning> [Accessed 18 Jan. 2015].
- Wiseman, A. (2013). Policy Responses to PISA in Comparative Education Perspective. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.
- Zhao, Y. and Meyer, H. D. (2013). High on PISA, Low on Entrepreneurship? What PISA Does Not Measure. In: H. D. Meyer and A. Benavot, eds. (2013). *PISA, power, and policy: the emergence of global educational governance*. Oxford: Symposium Books Ltd.