

Historical Roots of the Global Testing Culture in Education

Christian Ydesen and Karen Egedal Andreasen

Aalborg University, Denmark

Contact corresponding author: cyd@hum.aau.dk

ABSTRACT

Contemporary education is characterised by a global testing culture, reflecting the fact that students' learning outcomes and standards are the focus of policymakers worldwide. It therefore plays a significant role in educational policies in different national contexts. We offer a brief outline of the precursors and preconditions that have facilitated the rise of today's global testing culture. The article notes two chronological stages: the first encompasses a confluence of comparative education, the rise of applied psychology, and the formation of transnational organisational structures prior to World War II. The second stage features the emergence of international organisations immediately after World War II. We argue that these developments subsequently conflated into a trajectory fostered by Cold War policies and became dominant from the 1990s onwards.

Keywords: *testing culture, education, international organisations, history of education*

Significance of the Global Testing Culture and Its Antecedents

Given contemporary education, it is reasonable to speak of a global education space characterised by national education systems permeated by many similar components, such as marketisation, the greater use of tests and statistics, accountability requirements, international comparisons, and the mantra of raising standards (Plum, 2014; Smith, 2016). As Nordin and Sundberg (2014, p. 13f.) argue, 'today, making major reforms in the education sector without reference to global or transnational indicators seems politically stillborn'. A key constituent of this development has been the production of seemingly objective indicators and data deriving from international large-scale assessments.

Manuscript received: 12.12.2018. Manuscript accepted: 09.04.2019.

©2020 C. Ydesen & K. E. Andreasen. This is an Open Access article distributed under the terms of the Creative Commons CC-BY 4.0 License. ISSN 1891-5949, <https://noredstudies.org>

Citation: Ydesen, C. & Andreasen, K. E. (2020). Historical Roots of the Global Testing Culture in Education. *Nordic Studies in Education*, 40(2), 149–166. <https://doi.org/10.23865/nse.v40.2229>

Speaking of a global testing culture reflects the fact that students' learning outcomes and standards are the focus of policymakers (Addey et al., 2017; Hill & Kumar, 2009). Educational reforms have installed systems where performance measurements and test results are the main tools in quality assessments and the basis for parents' school choices; school funding; student, teacher, and school rankings; and school leader performance payments. As Smith (2016, p. 7) notes, 'the reinforcing nature of the global testing culture leads to an environment where testing becomes synonymous with accountability which becomes synonymous with education quality'. In other words, the conception of a global testing culture reflects the observation that practices in which rankings, performance indicators, and accountability based on various test results are in evidence across the globe (Rizvi & Lingard, 2010; Lindblad et al., 2015). In parallel, the global testing culture is closely affiliated with what Pasi Sahlberg has called the Global Education Reform movement (GERM). The GERM is an education reform approach that broadly follows the tenets of New Public Management and Neoliberalism. It is structured around a common set of policy ideas including standards-based management, performance evaluation, and accountability (Fuller & Stevenson, 2019).

Education data and their presentation frame and shape the political and public discourse on education: '[International large scale assessments] can be seen as a practice showing what is educationally possible' (Lindblad, Pettersson, & Popkewitz, 2015, p. 39). Furthermore, such assessments also influence the very ideas and ideals, purposes, values, and aims of schooling and teaching (Biesta, 2015). Our concern in this respect is not that national education systems are becoming uniform but that the data not only depict certain empirical findings but also express a normative worldview, which then is embodied in the very system of indicators (Desrosières, 1998; Rose, 1999). The testing culture thus ultimately affects educational access and social mobility, along with the performance of and benefits given to different groups. It also plays a significant role in educational policies and conditions in different national contexts (Allan & Artiles, 2017).

This testing culture, however, did not emerge *ex nihilo*. Its history features a long – but not necessarily coherent – development that can be traced to comparative education's foundation as a research field (Brickman, 1966, 2010), the establishment of international networks and organisations engaged with the field of education (Fuchs, 2007; Lawn, 2008), and the ascent of applied psychology in general and psychometrics in particular (Danziger, 1998). We argue that these precedents – building on an amalgam established mainly during the interwar years – conflated into a unified trajectory fostered by Cold War policies and grew dominant from the 1990s onwards.

To sustain this argument, we briefly outline the precursors, antecedents, and preconditions that facilitated the rise of today's global testing culture in education. The article considers two chronological stages: the first encompasses a confluence of comparative education, the rise of applied psychology, and the transnational organisational structures that began materialising prior to World War II (WWII). The second

stage features the emergence of foundations and organisations immediately after WWII, a period concerned with educational measurement and comparisons.

State of the Art: Sharpening our Focus

Significant policy research has been conducted on the functioning of the global testing culture (e.g. Grek, 2009; Meyer & Benavot, 2013; Rubenson, 2008; Smith, 2016). A key insight is that there has been no inevitable policy convergence due to international large-scale assessments. Instead, specific contextual factors seem to influence how test results and policy recommendations are interpreted and adapted for specific national schooling systems (Bieber & Martens, 2011; Carvalho & Costa, 2014). Conversely, comprehensive research (e.g. Grek, 2010; Lawn, 2011; Ozga et al., 2011) argues that experts and international organisations create data that transcends national policy debates, because the data enable cultural exchanges across borders and places, creating a new type of virtual, borderless policy space. This is a core feature of the global testing culture.

While policy studies examine the global testing culture's comparative impact, historical studies investigate its various components. American historiographers have explored how the foundations of contemporary educational testing rest on 19th-century developments (Reese, 2013). A main point of Reese (2013, p. 4) is that educational reformers prior to the American Civil War in 1861 'were the first to rank urban teachers, students, and schools based on quantitative scores, to shame the worst and honor the best'.

The majority of historical studies of educational testing are, as Reese's, tied to a national reference frame, mostly concerned with the North American context. The international – and even transnational – nature of the global testing culture has been addressed in only limited publications. Cardoso and Steiner-Khamsi's (2017) groundbreaking article examines education indicator research during three time periods. Their article is organised around three influential persons and institutions in the history of education indicator research: Jullien de Paris (1775–1848), Teachers College at Columbia University, and the United Nations Educational, Scientific and Cultural Organization (UNESCO) Institute for Statistics. Using these three focus points, the article finds discursive shifts in the policy usage of educational statistics affiliated with the three historical processes of modernisation/nation building, colonisation/development, and standardisation/globalisation. Their article thus describes a core development in comparative education. Although our article lies in the wake of Cardoso and Steiner-Khamsi, it offers a slightly different perspective on the necessary conditions – or core building blocks – of the contemporary global testing culture and adds other factors, such as applied psychology and the organisational landscape in the two chronological stages treated here.

In this regard, Lawn's 2008 volume concerning the International Examinations Inquiry (IEI) and 2014 followup article are pivotal. The IEI originated from a 1930s initiative of Columbia University and the Carnegie Corporation. Its purpose was to improve ways of identifying students suitable for secondary school (Hegarty, 2014). Apart from its focus on examinations, the inquiry also focused on intelligence, one of

the most important psychological issues at the time. Lawn (2014, p. 24) demonstrates that the IEI published what could be considered the first data-driven research inquiry into comparative education in nine countries and argues that the IEI formed 'a space in which pupil tests and statistical foundations prefigured the post-war expansion of comparative data in education and its use in governing education'. Lawn (p. 21) also notes that data collection on education began accelerating from the 1930s onwards: 'The growth of cross-border expert engagement in the mid-twentieth century created the basis for the later internationalization of education data and comparison'.

Lawn's work reflects the spatial turn in the history of education (Fuchs, 2014; Popkewitz, 2013), emphasising the importance of a transnational flow of expertise in the workings of global education. Its key components are the move beyond methodological nationalism and an understanding of the dynamics between place and space (Christensen & Ydesen, 2015; Lawn, 2014). As Nordin and Sundberg (2014, p. 15) state, education is 'transnational and national at the same time', meaning that place is understood as the setting or location, while space is where interaction, confluence, and exchanges happen.

Another relevant research field is the history of international organisations, such as the International Bureau of Education (IBE; e.g. Hofstetter & Schneuwly, 2013), UNESCO (e.g. Duedahl, 2016; Kulnazarova & Ydesen, 2016), the Organisation for Economic and Co-operative Development (OECD; e.g. Bürgi, 2016; Bürgi & Tröhler, 2018), and the International Association for the Evaluation of Educational Achievement (IEA; e.g. Landahl, 2017; Pizmony-Levy, 2013; Purves, 1987; Wagemaker, 2011). While this research certainly transcends national reference frames and offers interesting findings relating to the workings and impacts of these organisations and their roles in education, it is largely limited to specific decades. This observation also holds for Lawn's studies. However, we determine from these – and Barnett and Finnemore (2004) – that international organisations have great autonomy and significant power in shaping education globally.

Given these historiographies, this article offers a longterm perspective on 20th-century education history to enhance our understanding of the rise of the global testing culture. Although the article paints with a broad brush, the analysis contributes knowledge about recurring themes, perspectives, continuities, and ruptures in the history of the global testing culture in education.

Antecedents of the Global Testing Culture: Before WWII

This section focuses on the confluence of comparative education, the rise of applied psychology, and the organisational structures that began forming before WWII.

Comparative Education

The rise of comparative education as an academic field has a long history and constitutes a necessary condition for the contemporary global testing culture, even though comparativists could consider it an inadvertent development.

There are three main reasons for this connection. First, from its outset, comparative education instituted a comparative mindset – a logic based on the measurement, qualitative or quantitative, of one education system against another – with the aim of learning from comparisons to improve a given system (Cardoso & Steiner-Khamsi, 2017). From a historical perspective, the roots of such comparative studies can be traced far back, such as this early example from Friedrich August Hecht's 1795 book *De re scholastica anglia cum germanica comparata*, which compares schools in England with those in German states (Pettersen et al., 2015). Hecht succinctly expresses the later famous quotation of Sir Michael Ernest Sadler (1861–1943): 'What can we learn from the study of foreign systems?' (Bereday, 1964). Comparative education has also manifested itself in other practices, such as exhibitions and fairs, which became recurring events in the second half of the 19th century (Lundahl, 2016; Lundahl & Lawn, 2015). These exhibitions promoted the application of a comparative logic among national education systems and, as Sobe and Boven (2014) argue, 'international expositions allowed for educational systems and practices to be "audited" by lay and expert audiences'. Remember that, in the 19th century, the words *examination* and *exhibition* were often used synonymously (Reese, 2013, p. 2).

Second, comparative education was historically permeated by a distinct colonial discourse rooted in civilisation theory. It is therefore a Eurocentric approach to education, with a global outlook aimed at elevating the Third World. This strand in comparative education is still generally evident in the Programme for International Student Assessment (PISA) and the offshoot PISA for Development, both designed according to standards defined by the Global North (Cardoso & Steiner-Khamsi, 2017). Teachers College, Columbia University was a central hub for the expansion of American colonialism in education (Takayama, Sriprakash, & Connell, 2017). The point is that comparative education has often operated with hierarchisations in education systems and with varied notions about the best working practices in education.

Third, comparative education has been concerned with developing and refining an arsenal of methodologies and vocabularies for scientific and valid comparisons among education systems (Beech, 2006; Schriewer, 2012; Steiner-Khamsi, 2002) – note the concepts of juxtaposition, *tertium comparationis*, decontextualisation, borrowing, silent borrowing, and transferring, as well as the entire array of quantitative and statistical tools assembled to measure and sanctify the results (Bereday, 1967). As Cardoso and Steiner-Khamsi (2017, p. 401) state, 'the use of indicators makes educational systems comparable regardless of how different they are'.

Applied Psychology

Science and cooperation among its practitioners in different national contexts represent another backdrop for the birth of comparative practice within education. This pertains especially to psychology as both a science and a scientific field, which, since its earliest days, has been characterised by transnational cooperation and inspiration and the exchange of research results and theories (e.g. Hearnshaw, 1979).

Interestingly, educational testing appeared on the educational scene in most Western countries around the same time. Intelligence testing, for example, originated in Paris and travelled to California, Hamburg, New York, London, Edinburgh, and the rest of the world (Ydesen, 2011). Scientific standardisation was essential to this movement, since it enabled people to work across borders (Grek et al., 2009).

Due to the endeavours of psychologists to have psychology recognised and established as a 'real science' and academic field, some practitioners in this area adapted themselves to and were strongly influenced by the positivist paradigm dominating the late 1800s and early 1900s. Several education scholars committed themselves to research following positivist ideas – for instance, conducting controlled experiments or different tests to compare the results – and characterised by attempts to identify what could be considered general human traits, such as intelligence (Danziger, 1998). For such purposes, standardised testing was developed and soon became common as both a tool and a technology.

These ideas and trends went on to influence the realm of education. They entered this field through applied psychology and psychologists' questions related to education, for instance, as in experimental pedagogy – which had been founded circa 1900 (e.g. Claparède, 1911) – along with research on intelligence during the same period. These new theories and ideas were soon disseminated via publications and activities in associations and organisations, inspiring the practices of pedagogues, educational psychologists, and other professionals and academics throughout most of the Western world.

The rise of applied psychology in the interwar years was closely affiliated with the progressive educational movement. Many leading testing protagonists were members of and worked actively in such progressive education organisations as the New Education Fellowship (NEF; Ydesen, 2011). The progressive education movement at large was the standardbearer of a humanistic line of thought aimed at emancipating the child from its surrounding society, allowing it to develop freely. Conversely, testing protagonists were stimulated by an experimental scientific line aimed at disclosing the nature of the child and accommodating the educational system according to these findings to maximise society's perceived benefits. The common denominator between the wider audience of progressive educators and the testing protagonists was a critical attitude towards teachers' traditional examinations, which they considered a subjective evaluation tool, and an optimistic view of testing as a just and efficient differentiation tool compatible with meritocratic ideals (ibid.). Nonetheless, the testing protagonists tended to view pedagogy as merely applied psychology. Today, in the global testing culture, we are witnessing a similar reductionistic mechanism, in that education is transformed into learning and learning goals, given that learning is transformed into measurable performance according to such goals, while measurable performance is transformed into testing.

Organisational Structures

In terms of organisational structures, the NEF formed a space in which new progressive ideas could flourish, including notions about the benefits of mental tests. In

August 1929, the NEF held its largest conference in Denmark, with around 2,000 participants from 43 nations (Fuchs, 2004). The conference was very important in the international educational field and its report states, 'It is no exaggeration to say that this book contains the truest account available anywhere of the various currents of progressive educational thought in the world at this critical time' (Sadler, 1930, p. xi). A remarkable feature of the NEF conference was the first-time inclusion of a conference group titled 'Mental Tests' (Ydesen, 2011, p. 83).

The IBE also constitutes an interesting organisation. Drawing on the work of Rasmussen (2001), Hofstetter and Schneuwly (2013) argue that the IBE represents the transnational turn in the early 20th century. The IBE assigned itself the task of creating a platform to rally the numerous organisations at work worldwide that promoted intellectual cooperation, international solidarity, and educational renewal. Comparative education was upheld as the model discipline and its purpose was to 'bring together diversity and not to reduce it to unity' (ibid., p. 225).

These transnational organisations significantly promoted and inspired work with educational experimentation and crossborder initiatives. Numerous experiments were conducted across educational systems during the interwar and postwar periods, promoting a comparison mindset, even for those working in classroom settings.

Internationalisation of Education after WWII

The internationalisation of education prior to WWII was supported by different kinds of transnational cooperation among, for instance, scientists (e.g. psychologists), educationalists, and politicians. Formal associations mediated some of this cooperation, such as those focused on experimental pedagogy or the IBE. Different joint activities also served as mediators, such as exhibitions and psychological researchers' engagement with scholarship on intelligence and similar activities.

After WWII, these processes of sharing and disseminating knowledge internationally were influenced by the strengthening and formalisation of international cooperation in associations and organisations that were directly or indirectly addressing the educational sphere. This section briefly examines the paradigms and approaches to education prevalent among organisations such as UNESCO, founded in 1945; the IEA, which began operations in 1958; and the Organisation for European Economic Co-operation (OEEC)/OECD.¹

Balancing Education Ideals: Education for All, Effectiveness, and the Economy

Historically, UNESCO has embodied different ideals about education, ranging from education for peace and education for all to concerns about effective education

1 The World Bank has also played a role in shaping a global education space (Heyneman, 2003; Jones, 1992). For our purposes, however, we find that its economic approach to education is broadly covered by our discussion of other organisations.

planning and the improvement of countries' economies. Thus, UNESCO represents myriad points of view, not all necessarily compatible, with some rooted in pedagogical ideals and the universal purposes of the UN system, while others have pursued comparative perspectives based on the development of valid quantitative indicators.

To support the improvement of its member countries' educational systems, UNESCO realised early on the need for more systematic – and comparable – data for policymakers' educational planning and activities. At the Fourth UNESCO General Conference in 1949, a clearinghouse service was established, meant to provide member countries with different kinds of comparative information about national education, such as statistics and student performance assessments. Resolutions from the conference contain statements about a general education clearinghouse that read, 'The DirectorGeneral is instructed to maintain a clearing house in education' and 'To this end he shall: Arrange for educational missions to Member States, at their request and with their financial cooperation, for the purpose of making surveys, advising, and assisting in educational improvement, particularly in wardevastated or less developed regions' (UNESCO, 1949, p. 14).

In the 1950s, the systematic collection of educational statistics was thus seen as an activity UNESCO could manage and that generally and severally supported the collection of information about education systems, schools, and outcomes, including student performance. Additionally, the use of standardised testing played a central role in supporting data collection of a presumed comparative nature (Smyth, 2005). UNESCO (1949, p. 14) had a robust interest in the development of compulsory education systems and one of the first tasks assigned to the clearinghouse in 1950, in cooperation with the IBE, was to launch a study concerning 'problems involved in making free compulsory primary education more nearly universal and of longer duration throughout the world'.

In 1952, the UNESCO Institute for Education, originally focusing on comparative education, was founded (Elfert, 2015; Landsheere, 1997). Several conferences were held under its auspices during the 1950s. The institute hosted meetings for educational researchers where participants discussed such matters as measurement in education in general, evaluation, and problems related to examinations in educational systems. The meetings were attended by prominent researchers then dominating the field, such as Swedish psychologist Torsten Husén and American educational psychologist Benjamin Bloom. The attendees shared an interest in crossnational – and thus comparative – research within education and attempted to use comparative research to address various educational problems. For instance, individual countries were considered too small and homogeneous to explain differences in school performance (Landahl, 2017). These meetings nurtured ideas on how to conduct large comparative international surveys, the first attempt initiated in the late 1950s with a pilot study called the 'Twelve Country Study' (Keeves, 2011; Landsheere, 1997). The project was successful and formation of the IEA was initiated soon after, with, among others, Husén and Danish psychometrician Georg Rasch as important contributors (Keeves, 2011).

In 1964, UNESCO's International Institute for Educational Planning was founded. The 1960s marked a crucial period in establishing a new economic paradigm in the approach to education, drawing on manpower planning and human capital theory. The OECD's Mediterranean Regional Project represented this new era in educational planning. At the same time, however, the more general interest in pedagogy and improving teaching that dominated part of educational research since the late 1800s – for instance, in the progressive pedagogy movement and experimental pedagogy – took on new directions. Such interest merged with new assessment technologies and statistical methods and supported new educational research practices focusing on developing what was considered an evidence-based and efficient pedagogy – efficient, that is, in the sense of pedagogy leading to strong subject-specific test performance.

Seeking an Evidence-Based and Efficient Pedagogy

The interest in improving pedagogy and supporting efficiency in education soon created a new and dominating practice within some areas of educational research. Researchers from different scientific areas – such as educational psychology, comparative education, intelligence testing, and the statistics of education – found a common interest in attempts to improve basic teaching and students' performance. The new technologies to assess and conduct surveys facilitated the collection, analysis, and comparison of large datasets across national education systems. These early international largescale assessments were also important tools paving the way for new attempts to improve educational systems and identify so-called 'best practices' and efficient pedagogy understood and identified based on test results.

The IEA was formed under the influence of such interests and trends. However, in employing new technologies, this research became dominated by positivistically inspired approaches and practices based on the collection and comparison of quantitative data in large-scale and often international surveys. Efficiency and quality in pedagogy were thus identified by different performance measures and economic factors and best practices became what appeared to be economically the most affordable and rational practices in light of such performance measures. Through the impact of these processes and results, the educational sphere became influenced by concerns other than pedagogy, didactics, educational ideals, and nation building, all of which had hitherto dominated education in many countries.

Since its founding, the IEA has conducted numerous international educational comparative surveys and studies, such as the Six Subject Study in 1966–1973 (Walker, 1976), the Trends in International Mathematics and Science Study (TIMMS), the Progress in International Reading Literacy Study (PIRLS), and other surveys playing important roles in educational policies today (Pizmony-Levy, 2013). The results of the first PISA round published in 2001 were central in the political decisions and processes leading to the standardized national testing programme in the Danish public school in 2005/2006. But the results of an international IEA survey on student reading

performance conducted in the late 1980s, showing that Danish pupils did not perform as well as pupils in some of the other Nordic countries, can be seen as forming the early background for the implementation of such a testing practice, by changing the predominant understanding of Danish pupils as being skilled readers (Andreasen, Kelly, Kousholt, McNess, & Ydesen, 2015; Gustafsson, 2012).

Another addition to these comparative endeavours was the school effectiveness movement appearing in the late 1970s that focused on 'effective schools' and worked to identify best practices in pedagogy and school leadership. The movement can be viewed as paralleling the IEA, since it was based on similar ideas (Goldstein & Woodhouse, 2000; Townsend, 2007). The movement manifested itself as a formal organisation in 1988, with the International Congress for School Effectiveness and Improvement, which published a journal and convened an annual congress. Its focus has been on identifying 'effective teaching and leadership' using a variety of international surveys. The movement has gained a strong footing in some countries via such reports as 'Exceptional Effectiveness: Taking a Comparative Perspective on Educational Performance' (Harris & Hargreaves, 2015). The IEA and the school effectiveness movement can be categorised as the promoters of an influential *what works–best practice–evidence based policy paradigm* popular in contemporary education policy (Connell, 2013). Thus, a picture emerges of certain international organisations serving as arbiters of a positivist statistical agenda in education policy.

UNESCO's reasons for launching new initiatives were largely informed by its aims to expand and strengthen compulsory education for purposes aligned with offering development, extending modern citizens' skills, and promoting international understanding (Boel, 2016). Yet another player would enter the educational arena in the 1960s in support of the *what works* paradigm noted above: the OECD, a highly influential organisation that also heavily promoted international comparisons across national school systems.

For decades, the OECD has promoted a vision of education as one of providing human capital to improve the economies of nationstates (Papadopoulos, 2011; Tröhler, 2010). While the OECD is essentially an economic organisation, education appeared on the OEEC agenda in 1958 due to the Soviet Sputnik satellite launch the previous year (Kogan, 1979; Tröhler, 2010). Education gradually came to play a defining role in understanding the economic capabilities and potential of nationstates (Petterson, 2014; Ydesen, 2013). Since then, the OECD has developed into one of the most powerful agencies in terms of shaping a global education space, because of its country reviews, test programmes, and reports (Bürgi, 2012; Grek, 2009; Martens, 2007; Moutsios, 2009).

In 1961, the first OECD conference on education was held in Washington, DC. It is indeed telling that one of its key speakers opined, 'May I say that, in this context, the fight for education is too important to be left solely to the educators' (OECD, 1961, p. 35). Education was becoming increasingly politicised, having transformed into a battlefield in the context of the Cold War.

In 1962, the Programme for Educational Investment and Planning (EIP) was launched. Among other things, it called for member countries to gather comprehensive statistical data. The next year, the OECD request prompted the Danish Ministry of Education to hire an economics and statistical counsellor. Besides providing data to the OECD, for example, on teacher–student ratios, factors affecting student choice in education programmes, and progress reports on educational investment planning, the counsellor was tasked with advising central and local authorities about educational investment planning.² The EIP considers that education must employ more effective planning processes using the latest quantitative methods to optimise its results regarding economic growth and thus win the technology race against the Eastern bloc.

In 1968, to strengthen its focus and initiatives concerning educational improvements, the OECD founded the Centre for Educational Research and Innovation (CERI). Jarl Bengtsson (2008, p. 1), former head of CERI, notes that a feature of the centre's formation was 'the emergence of education as a nascent field of research and analysis at a time of rising investments and expectations for education'. Thus, CERI was established during a period when the role of education in the democratic welfare states had become obvious and the centre was explicitly created for policy research (Schuller, 2005). The OECD (2016, p. 18) describes CERI's purpose by explaining, 'a large body of CERI work has been founded on the need for educational decisionmaking to be better informed by evidence, by awareness of what is taking place in other countries'. The OECD has since constructed a huge database of statistical figures on both member and nonmember countries in the field of education.

Conclusion

The global testing culture dominating current educational policies and practices worldwide has a lengthy and fascinating pedigree, as we described. The historical developments presented represent a necessary but not sufficient conditions for the rise of the global testing culture, that is, they should be considered stepping stones for the contemporary workings of global education. The processes leading to the global testing culture's formation include developments and practices from numerous scientific and political areas. Some seem to have merged over time, even given different origins, along with differing and even conflicting purposes at points.

The years before WWII witnessed the first steps in the formation of a new comparative practice in educational research. Inspired by ideas from experimental pedagogy and developments within psychology – including the rise of mental testing – and driven by efforts to improve educational systems as well as a common and more general interest in educational research, such initiatives gained a new platform and were

2 Danish National Archives, Ministry of Education, International Office, 1959–1970 Cases Concerning International Organisations, OE 2 1963 – 4 1963, General Memorandum, 9 November 1964, Working Programme for the EIP team, Ministry of Education in Denmark, p. 3.

made possible by extended transnational cooperation in the West. However, while the initiatives of the IEI, among others, seen in this context, were influenced and temporarily halted by WWII, the end of WWII and the post-war era marked the beginning of a new era of transnational cooperation in newly established organisations such as UNESCO. Furthermore, two tendencies seemed to merge and form a new practice in this context: on the one hand, researchers' interest in improving pedagogy and, on the other, politicians' and economists' interest in making educational systems more economically efficient.

The process has been dominated by organisations such as UNESCO, the IEA, and the OECD, even though they have supported such activities for differing reasons and purposes, with UNESCO and the IEA focusing on improving pedagogy and identifying best practices, in contrast to the OECD, which pursues a clearly defined economic policy agenda.

Before the 1990s, international comparative assessments in education were primarily initiated and administered by such nongovernmental organisations as the IEA; however, since the 1990s, the OECD also adapted and launched such assessments. The OECD's well-established authority conveys high status in member as well as nonmember countries, which strengthens the impact of both the processes and results.

The comparative turn in global education policy advocated and promoted by the OECD must be understood in light of crossnational comparison being considered the best engine to promote educational quality (Martens, 2007). Note, however, that this observation entails a shift from research to policy (Wagemaker, 2013), as well as a shift in focus from pedagogic practice to academic performance. In other words, the OECD has pursued a path of identifying best practices designed to improve education systems around the world by using comparisons and through the development of various monitoring tools. This activity has often been accomplished in close conjunction with the European Commission engaged in the mutual identification of educational problems (Grek, 2010).

The global testing culture has been strongly criticised for its influence on school systems and pedagogy. Its core features are stronger emphasis on national and international comparisons, student performance, and the control of education – for instance, learning goals and corresponding assessments and standardised testing at the national level. These methods have been criticised for sacrificing a focus on pedagogy and *Bildung*, whose success is more difficult to assess (Biesta, 2015). In addition, the global testing culture tends to strongly influence what is considered normal and leaves less room for deviations therefrom. Consequently, cultural and/or language minorities are at risk of discrimination in these processes (Andreasen & Kousholt, 2018).

Recently, critical voices have spoken out against not only these processes but also the *organisations* orchestrating them – the OECD, PISA, the IEA – and their political influence in member and even non-member states.

One point of criticism addresses the data and information generated and distributed: the underlying conditions of statistics are difficult to determine. Even though

skilled educational statisticians have strongly criticised conclusions drawn from the data, they seem to have little influence (e.g. Kreiner & Christensen, 2014). Another point of contention highlights the conflict between democratic ideals and governance guided by comparative statistics. Organisations such as the OECD are political by nature but their influence on education in both member and non-member states has become increasingly direct (Lewis, 2017). Such direct influence compromises and threatens democracy and democratic processes but explains the recent uniform developments of educational systems. For instance, representatives to PISA's governing board are *appointed* by each member country (OECD, 2017), such that individuals serving in such a capacity are not democratically accountable. Such problematics could not have been predicted at the outset of these processes but, given their gravity, they must be paid careful attention in the future.

REFERENCES

- Addey, C., Sellar, S., Steiner-Khamsi, G., Lingard, B., & Verger, A. (2017). The rise of international large-scale assessments and rationales for participation. *Compare: A Journal of Comparative and International Education*, 47(3), 434–452. <https://doi.org/10.1080/03057925.2017.1301399>
- Allan, J., & Artiles, A. J. (Eds.). (2017). *World yearbook of education 2017: Assessment inequalities*. London; New York: Routledge, Taylor & Francis Group.
- Andreasen, K. E., & Ydesen, C. (2016). School accountability, educational performance testing and inequalities in a global perspective 1945 to present. Abstract from The Road to Global Inequity, Aarhus, Denmark.
- Andreasen, K. E., Kelly, P., Kousholt, K., McNess, E., & Ydesen, C. (2015). Standardised testing in compulsory schooling in England and Denmark: A comparative study and analysis. *Bildung und Erziehung*, 68(3), 329–348.
- Andreasen, K. E. & Kousholt, K. (2018). Minorities and education testing in schools in arctic regions: An analysis and discussion focusing on normality, democracy, and inclusion for the cases of Greenland and the Swedish Sami schools. In B. Hamre, A. Morin & C. Ydesen (Eds.), *Testing and inclusive schooling – international challenges and opportunities* (pp. 19–33). London: Routledge.
- Barnett, M. N., & Finnemore, M. (2004). *Rules for the world: International organizations in global politics*. Ithaca, NY: Cornell University Press.
- Beech, J. (2006). The theme of educational transfer in comparative education: A view over time. *Research in Comparative and International Education*, 1(1), 2. <https://doi.org/10.2304/rcie.2006.1.1.2>
- Bengtsson, J. (2008). *OECD's Centre for educational research and innovation – 1968 to 2008*. Paris: Centre for Educational Research and Innovation.
- Bereday, G. Z. F. (1964). Sir Michael Sadler's 'Study of foreign systems of education'. *Comparative Education Review*, 7(3), 307–314.
- Bereday, G. Z. F. (1967). Reflections on comparative methodology in education, 1964–1966. *Comparative Education*, 3(3), 169–287. <https://doi.org/10.1080/0305006670030304>
- Bieber, T., & Martens, K. (2011). The OECD PISA study as a soft power in education? Lessons from Switzerland and the US. *European Journal of Education*, 46(1), 101–116.
- Biesta, G. J. J. (2015). Resisting the seduction of the global education measurement industry: Notes on the social psychology of PISA. *Ethics and Education*, 10(3), 348–360.

- Boel, J. (2016). UNESCO's fundamental education program, 1946–1958: Vision, actions and impact. In P. Duedahl (Ed.), *A history of UNESCO: Global actions and impacts*. London: Palgrave Macmillan.
- Brickman, W. W. (1966). Prehistory of comparative education to the end of the eighteenth century. *Comparative Education Review*, 10(1), 30–47.
- Brickman, W. W. (2010). Comparative education in the nineteenth century. *European Education*, 42(2), 46–56. <https://doi.org/10.2753/EUE1056-4934420206>
- Bu, L. (1997). International activism and comparative education: Pioneering efforts of the International Institute of Teachers College, Columbia University. *Comparative Education Review*, 41(November), 413–434.
- Bürgi, R. (2012). Bypassing federal education policies. The OECD and the case of Switzerland. *International Journal for the Historiography of Education*, 1(2), 24–35.
- Bürgi, R. (2016). Systemic management of schools: The OECD's professionalisation and dissemination of output governance in the 1960s. *Paedagogica Historica*, 52(4), 408–422. <https://doi.org/10.1080/00309230.2016.1178780>
- Bürgi, R., & Tröhler, D. (2018). Producing the 'right kind of people': The OECD education indicators in the 1960s. In S. Lindblad, D. Petterson & T. S. Popkewitz (Eds.), *Numbers, education and the making of society: International assessments and its expertise*. New York: Routledge.
- Cardoso, M., & Steiner-Khamsi, G. (2017). The making of comparability: Education indicator research from Jullien de Paris to the 2030 sustainable development goals. *Compare: A journal of comparative and international education*, 47(3), 388–405. <https://doi.org/10.1080/03057925.2017.1302318>
- Carvalho, L. M., & Costa, E. (2014). Seeing education with one's own eyes and through PISA lenses: Considerations of the reception of PISA in European countries. *Discourse: Studies in the Cultural Politics of Education*, 36(5), 638–646, DOI 10.1080/01596306.2013.871449.
- Christensen, I. L., & Ydesen, C. (2015). Routes of knowledge: Toward a methodological framework for tracing the historical impact of international organizations. *European Education*, 47(3), 274–288. <https://doi.org/10.1080/10564934.2015.1065392>
- Claparède, E. (1911). *Experimental pedagogy and the psychology of the child*. New York: Longmans, Green and Co., London: Edward Arnold.
- Connell, R. (2013). The neoliberal cascade and education: An essay on the market agenda and its consequences. *Critical Studies in Education*, 54(2), 99–112. <https://doi.org/10.1080/17508487.2013.776990>
- Danziger, K. (1998). *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press.
- Depaepe, M., & Smeyers, P. (2008). Educationalization as an ongoing modernization process. *Educational Theory*, 58(4), 379–389. <https://doi.org/10.1111/j.1741-5446.2008.00295.x>
- Desrosières, A. (1998). *The politics of large numbers: A history of statistical reasoning*. Cambridge, MA: Harvard University Press.
- Dorn, S., & Ydesen, C. (2014). Towards a comparative and international history of school testing and accountability. *Education Policy Analysis Archives*, 22, 115.
- Duedahl, P. (2016). *A history of UNESCO: Global actions and impacts*. London: Palgrave Macmillan.
- Elfert, M. (2013). Six decades of educational multilateralism in a globalising world: The history of the UNESCO Institute in Hamburg. *International Review of Education*, 59(2), 263–287. <https://doi.org/10.1007/s11159-013-9361-5>
- Fuchs, E. (2004). Educational sciences, morality and politics: International educational congresses in the early twentieth century. *Paedagogica Historica*, 40(5), 757–784.

Historical Roots of the Global Testing Culture in Education

- Fuchs, E. (2007). The creation of new international networks in education: The League of Nations and educational organizations in the 1920s. *Paedagogica Historica*, 43(2), 199–209. <https://doi.org/10.1080/00309230701248305>
- Fuchs, E. (2014). History of education beyond the nation? Trends in historical and educational scholarship. In B. Bagchi, E. Fuchs & K. Rousmaniere (Eds.), *Connecting histories of education – transnational and crosscultural exchanges on (post)colonial education* (pp. 11–26). New York: Berghahn Books.
- Fuller, K., & Stevenson, H. (2019). Global education reform: Understanding the movement. *Educational Review*, 71(1), 1–4. <https://doi.org/10.1080/00131911.2019.1532718>
- Goldstein, H., & Woodhouse, G. (2000). School effectiveness research and educational policy. *Oxford Review of Education*, 26(3/4), 353–363.
- Grek, S. (2009). Governing by numbers: The PISA effect in Europe. *Journal of Education Policy*, 24(1), 23–37.
- Grek, S. (2010). International organisations and the shared construction of policy ‘problems’: Problematisation and change in education governance in Europe. *European Educational Research Journal*, 9(3), 396. <https://doi.org/10.2304/eeerj.2010.9.3.396>
- Gustafsson, R. L. (2012). *What did you learn in school today? How ideas mattered for policy changes in Danish and Swedish schools 1990–2011*. Aarhus. Forlaget Politica.
- Harris, A., & Hargreaves, A. (2015). Exceptional effectiveness: Taking a comparative perspective on educational performance. International Congress for School Effectiveness and School Improvement (Monograph Series 3).
- Hearnshaw, L. S. (1979). *Cyril Burt, psychologist*. Ithaca, NY: Cornell University Press.
- Hegarty, S. (2014). From opinion to evidence in education: Torsten Husén’s contribution. In A. Nordin & D. Sundberg (Eds.), *Transnational policy flows in European education: The making and governing of knowledge in the education policy field* (pp. 21–32). Oxford: Symposium Books.
- Heyneman, S. P. (2003). The history and problems in the making of education policy at the World Bank 1960–2000. *International Journal of Educational Development*, 23(3), 315–337. [https://doi.org/10.1016/S0738-0593\(02\)00053-6](https://doi.org/10.1016/S0738-0593(02)00053-6)
- Hill, D., & Kumar, R. (Eds.) (2009). *Global neoliberalism and education and its consequences*. New York: Routledge.
- Hofstetter, R., & Schneuwly, B. (2013). The International Bureau of Education (1925–1968): A platform for designing a ‘chart of world aspirations for education’. *European Educational Research Journal*, 12(2), 215. <https://doi.org/10.2304/eeerj.2013.12.2.215>
- Jones, P. W. (1992). *World Bank financing of education. Lending, learning and development*. London and New York: Routledge.
- Keeves, J. (2011). IEA – From the beginning in 1958 to 1990. In C. Papanastasiou, T. Plomp & E. C. Papanastasiou (Eds.), *IEA 1958–2008: 50 years of experiences and memories* (pp. 3–40). Amsterdam: The International Association for the Evaluation of Educational Achievement.
- Kogan, M. (1979). *Education policies in perspective: An appraisal of OECD country educational policy reviews*. Paris: Organisation for Economic Cooperation and Development.
- Kreiner, S., & Christensen, K. B. (2014). Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika*, 79(2), 210–231. <https://doi.org/10.1007/s11336-013-9347-z>
- Kulnazarova, A., & Ydesen, C. (Eds.) (2016). *UNESCO without borders: Educational campaigns for international understanding*. New York, NY: Routledge.
- Landahl, J. (2017). Smallscale community, largescale assessment: IEA as a transnational network. Conference paper presented at the European Conference for Educational Research (ECER) in Copenhagen, Denmark, 22–25th of August.

- Landsheere, G. (1997). IEA and UNESCO: A history of working cooperation. Retrieved Nov. 21st, 2017 from <http://www.unesco.org/education/pdf/LANDSHEE.PDF>
- Lawn, M. (Ed.) (2008). *An Atlantic crossing? The work of the International Examination Inquiry, its researchers, methods, and influence*. Oxford, UK: Symposium Books.
- Lawn, M. (2011). Standardizing the European education policy space. *European Educational Research Journal*, 10, 259–72. <http://dx.doi.org/10.2304/eerj.2011.10.2.259>
- Lawn, M. (2014). Nordic connexions: Comparative education, Zilliacus and Husén, 1930–1960. In A. Nordin & D. Sundberg (Eds.), *Transnational policy flows in European education: The making and governing of knowledge in the education policy field* (pp. 21–32). Oxford: Symposium Books.
- Lewis, S. (2017). Policy, philanthropy and profit: The OECD's PISA for schools and new modes of hierarchical educational governance. *Comparative Education*, 1–20. <https://doi.org/10.1080/03050068.2017.1327246>
- Lindblad, S., Pettersson, D. & Popkewitz, T. S. (2015). International comparisons of school results: A systematic review of research on large scale assessments in education. A report from the *Educational Research Project SKOLFORSK*. Swedish Research Council. <https://doi.org/10.13140/RG.2.2.23176.01286>
- Lundahl, C. (2016). Swedish education exhibitions and aesthetic governing at world's fairs in the late nineteenth century. *Nordic Journal of Educational History*, 3(2), 3–30.
- Lundahl, C., & Lawn, M. (2015). The Swedish schoolhouse: A case study in transnational influences in education at the 1870s world fairs. *Paedagogica Historica*, 51(3), 319–334. <https://doi.org/10.1080/00309230.2014.941373>
- Martens, K. (2007). How to become an influential actor – the 'comparative turn' in OECD education policy. In K. Martens, A. Rusconi & K. Lutz (Eds.), *Transformations of the state and global governance* (pp. 40–56). London: Routledge.
- Meyer, H. D., & Benavot, A. (Eds.) (2013). *PISA; power, and policy: The emergence of global educational governance*. Oxford: Symposium Books.
- Mitch, D. (2005). Education and economic growth in historical perspective. *EH.Net Encyclopedia*, edited by Robert Whaples. July 26. URL <http://eh.net/encyclopedia/education-and-economic-growth-in-historical-perspective/>
- Moutsios, S. (2009). International organisations and transnational education policy. *Compare*, 39(4), 469–481.
- Nordin, A., & Sundberg, D. (Eds.) (2014). *Transnational policy flows in European education: The making and governing of knowledge in the education policy field*. Oxford: Symposium Books.
- OECD (1961). *Policy conference on economic growth and investment in education. Washington 16th–20th October 1961*. Paris: OECD Publishing.
- OECD (2016). *Trends shaping education 2016*. Paris: OECD Publishing. http://dx.doi.org/10.1787/trends_edu-2016-en
- Ozga, J., DahlerLarsen, P., Segerholm, C., & Simola, H. (Eds.) (2011). *Fabricating quality in education: Data and governance in Europe*. London: Routledge.
- Papadopoulos, G. S. (2011). The OECD approach to education in retrospect: 1960–1990. *European Journal of Education*, 46(1), 85–86.
- Petterson, D. (2014). The development of the IEA: The rise of largescale testing. In A. Nordin & D. Sunderberg (Eds.), *Transnational policy flows in European education* (pp. 105–122). Oxford: Symposium Books.
- Pettersson, D., Popkewitz, T. S., & Lindblad, S. (2015). On the use of educational numbers: Comparative constructions of hierarchies by means of largescale assessments. *Espacio, Tiempo y Educación*, 3(1), 177–202. <https://doi.org/10.14516/ete.2016.003.001.10>

Historical Roots of the Global Testing Culture in Education

- Pizmony-Levy, O. (2013). Testing for all: The emergence and development of international assessment of student achievement, 1958–2012. PhD dissertation, Indiana University.
- Plum, M. (2014). A 'globalised' curriculum – international comparative practices and the preschool child as a site of economic optimisation. *Discourse: Studies in the Cultural Politics of Education*, 35(4), 570–583. <https://doi.org/10.1080/01596306.2013.871239>
- Popkewitz, T. S. (2013). *Rethinking the history of education: Transnational perspectives on its questions, methods, and knowledge*. New York: Palgrave Macmillan.
- Purves, A. C. (1987). The evolution of the IEA: A memoir. *Comparative Education Review*, 31(1), 10–28. <https://doi.org/10.1086/446653>
- Rasmussen, A. (2001). Tournant, inflexions, ruptures: le moment internationaliste. *Mille neuf cents. Revue d'histoire intellectuelle*, 19(1), 27–41. <http://www.cairn.info/revue-mil-neuf-cent-2001-1-page-27.htm>.
- Reese, W. J. (2013). *Testing wars in the public schools: A forgotten history*. Cambridge, MA: Harvard University Press.
- Rizvi, F., & Lingard, B. (2010). *Globalizing education policy*. London, New York: Routledge.
- Rose, N. (1999). Powers of freedom: Reframing political thought. Cambridge: Cambridge University Press.
- Rubenson, K. (2008). OECD educational policies and world hegemony. In R. Mahon & S. McBride (Eds.), *The OECD and transnational governance* (pp. 293–314). Vancouver: British Columbia University Press.
- Sadler, M. (1930). Introduction. In W. Boyd (Ed.), *Towards a new education – A record and synthesis of the discussions on the new psychology and the curriculum at the fifth World Conference of the New Education Fellowship held at Elsinore, Denmark, in August 1929* (pp. xi–xvii). London: A. A. Knopf.
- Schriewer, J. (Ed.) (2012). *Discourse formation in comparative education* (4th rev. ed.). Frankfurt am Main: Peter Lang.
- Smith, W. C. (Ed.) (2016). *The global testing culture: Shaping education policy, perceptions, and practice*. Oxford: Symposium Books.
- Smyth, J. A. (2005). UNESCO's international literacy statistics 1950–2000. Background paper prepared for the Education for All Global Monitoring Report 2006. Literacy for Life. United Nations Educational, Scientific, and Cultural Organization.
- Sobe, N. W., & Boven, D. T. (2014). Nineteenthcentury world's fairs as accountability systems: Scopic systems, audit practices and educational data. *Education Policy Analysis Archives*. <https://doi.org/10.14507/epaa.v22.1673>
- Steiner-Khamsi, G. (2002). Re-framing educational borrowing as a policy strategy. In M. Caruso & H.-E. Tenorth (Eds.), *Internationalisierung: Semantik und Bildungssystem in vergleichender Perspektive [Internationalisation: Comparative education systems and semantics]* (pp. 305–343). Frankfurt am Main: Peter Lang.
- Takayama, K., Sriprakash, A., & Connell, R. (2017). Toward a postcolonial comparative and international education. *Comparative Education Review*, 61(S1), S1–S24. <https://doi.org/10.1086/690455>
- Townsend, P. (2007). Preface. In P. Townsend (Ed.), *International handbook of school effectiveness and improvement: Part 1* (pp. 3–26). Springer International Handbooks of Education. Dordrecht: 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. <https://doi.org/10.1007/s11217-009-9155-1>

- Tröhler, D. (2015). The medicalization of current educational research and its effects on education policy and school reforms. *Discourse: Studies in the Cultural Politics of Education*, 36(5), 749–764. <https://doi.org/10.1080/01596306.2014.942957>
- United Nations Educational, Scientific, and Cultural Organization. (1949). Recommendations concerning the direction of school programs towards international peace and security. Resolution no. 2.513. Executive Board.
- Wagemaker, H. (2011). IEA: International studies, impact and transition. In C. Papanastasiou, T. Plomp & E. Papanastasiou (Eds.), *IEA 1958–2008: 50 years of experiences and memories* (pp. 253–272). Amsterdam: International Association for the Evaluation of Educational Achievement.
- Wagemaker, H. (2013). International largescale assessments: From research to policy. In L. Rutkowski, M. von Davier & D. Rutkowski (Eds.), *Handbook of international large scale assessment: Background, technical issues, and methods of data analysis* (pp. 11–36). Boca Raton, FL: CRC.
- Walker, D. A. (1976). *The IEA six subject survey: An empirical study of education in twenty-one countries*. Stockholm: Almqvist & Wiksell International, New York: J. Wiley.
- Ydesen, C. (2011). *The rise of highstakes educational testing in Denmark, 1920–1970*. Frankfurt am Main: Peter Lang.
- Ydesen, C. (2013). Educational testing as an accountability measure: Drawing on twentiethcentury Danish history of education experiences. *Paedagogica Historica*, 49(5), 716–733. <https://doi.org/10.1080/00309230.2013.815235>