



Compare: A Journal of Comparative and International Education

 **Routledge**
Taylor & Francis Group

ISSN: 0305-7925 (Print) 1469-3623 (Online) Journal homepage: <http://www.tandfonline.com/loi/ccom20>

The making of comparability: education indicator research from Jullien de Paris to the 2030 sustainable development goals

Manuel Cardoso & Gita Steiner-Khamsi

To cite this article: Manuel Cardoso & Gita Steiner-Khamsi (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, DOI: [10.1080/03057925.2017.1302318](https://doi.org/10.1080/03057925.2017.1302318)

To link to this article: <http://dx.doi.org/10.1080/03057925.2017.1302318>



Published online: 07 Apr 2017.



Submit your article to this journal [↗](#)



Article views: 4



View related articles [↗](#)



View Crossmark data [↗](#)

Full Terms & Conditions of access and use can be found at
<http://www.tandfonline.com/action/journalInformation?journalCode=ccom20>

Download by: [178.197.233.134]

Date: 10 April 2017, At: 03:07



The making of comparability: education indicator research from Jullien de Paris to the 2030 sustainable development goals

Manuel Cardoso^a  and Gita Steiner-Khamsi^b

^aUNICEF, New York, USA; ^bTeachers College, Columbia University & Graduate Institute of International and Development Studies, New York, USA

ABSTRACT

The authors examine indicator research over three periods and discuss shifts in policy usage over time. The study compares influential actors that reflect discursive shifts in how, and for what purpose, indicators were used: (1) Jullien de Paris, (2) faculty at Teachers College at Columbia University in the early-twentieth century (notably Paul Monroe and Isaac Kandel) and (3) UNESCO Institute for Statistics' contribution to the debate on the 2030 Sustainable Development Goals (SDGs) indicators. Arguably, these periods reflect the broader agenda and usage of indicators: modernization/nation-building (Jullien), colonization/development (Monroe and Kandel) and standardization/globalization (SDGs). In this view, indicators make systems comparable, despite their differences: educational systems are not comparable *per se*, they are made comparable through standardized measurement. Jullien crucially shaped empirical research through uniform questions that enable description, analysis and comparison educational systems' key features. Recently, policy analysts and researchers have referred to those standardised questions as indicators.

KEYWORDS

Indicators; history; comparability; development; globalization; assessment

1. Introduction

Marc-Antoine Jullien de Paris' (1775 – 1848) *Esquisse d'un ouvrage sur l'éducation comparée* (Plan and preliminary views for a work on comparative education) is widely acknowledged as the methodological foundation for comparing systems and producing statistics, as well as the first attempt to reflect on suitable indicators for measuring the provision and outcomes of education. In Jullien de Paris' surveys of national educational systems, he used the same set of questions to compare key features of educational systems such as, for example, school entrance age, duration of primary school and so on. (Jullien de Paris 1817). In addition to shaping the comparative method of cross-national inquiry, he helped to justify the need for comparison: lesson-drawing, policy borrowing or, better known today, learning from 'best practices' and international standards. The two contributions, indicator research and

policy learning, remain closely related and have helped build the foundation for comparative education as an applied professional field and academic area of research.

This article examines indicator research during three distinct time periods and discusses shifts in policy usage over time. We have selected influential persons and institutions of different types (Jullien de Paris; Teachers College at Columbia University; UNESCO Institute for Statistics [UIS]) for each of the three time periods that reflect discursive shifts that occurred in how, and for what purpose, indicators were used:

- Jullien de Paris' (1817) *Esquisse d'un ouvrage sur l'éducation comparée*
- Research at Teachers College, particularly Paul Monroe's 'surveys' of educational systems in the Philippines (1912–1914) and Isaac Kandel's work around comparability and Civilization Theory
- UIS's collaboration with other agencies and stakeholders in the development of education indicators for the 2030 Sustainable Development Goals (SDGs) (2015–2016)

Admittedly, these actors have different statuses and links to various spheres of power. Without a formal position in an education system, Jullien's power was limited, although he was part of an impressive informal network of researchers and practitioners. (Fraser 1964) Teachers College forged influential partnerships with the US government with a view to international projects. (Monroe, Páez, and Duggan 1925) UIS has a mandate to produce internationally comparable data in the areas of UNESCO's competence (Cussó 2006).

Nonetheless, they share a commonality: their work in education, in various geographic and organisational contexts, was produced at foundational moments in the broader field of official statistics. In France the bureau of statistics of the Republic was founded in 1800, closed by Napoleon in 1812 and replaced by the *Statistique Générale de la France* in 1833; in the USA, the Bureau of the Census became permanent in 1902 (Desrosières and Naish 1998). The UIS was founded in 1999, and moved away from UNESCO's headquarters in Paris in 2001 (Cussó 2006), coinciding with the Millennium Development Goals (MDGs) and the Education for All (EFA) goals.

In addition, the choice of actors is necessarily partial. For instance, UIS is an offshoot of UNESCO, which in 1946 succeeded the International Institute of International Cooperation (Smyth 2008); therefore, this thread would lead to the 1920s, Monroe's heyday. However, our aim here is not the history and genealogy of each actor, but proposing examples of different educational comparison approaches since Jullien's time.

Naturally, the uses *and* abuses of comparative policy studies have changed significantly over the past two hundred years. In concert with other scholars who explore the history of cross-national comparison and educational statistics (see, for example, Cussó and D'Amico 2005), we discern three discursive shifts in the policy usage of educational statistics: in the early period (reflected in Jullien de Paris' *Esquisse*) comparison was used for modernisation, viewed at the time as a feature of nation building. The explicit objective of cross-national comparison was to ensure that educational provisions and practices from nations that had already successfully modernised by introducing public education could be adopted by other nations seeking to modernise as well.

In the early-twentieth century – marking the second period of study in this article – cross-national comparison became part of colonisation and 'development'. The assumption that the educational systems of the colonial empires were superior to those of their colonies permeated the social sciences and humanities. As a result of the sweeping de-colonisation

movement occurring by the mid-twentieth century, the colonisation project was replaced with ‘aid’ in the form of ‘technical assistance’ or ‘international cooperation’.

The third period under investigation began in the 1990s and reflects an agenda of global trade relying heavily upon the transnational flow of people, production, finance, ideas, beliefs and technology. Today, the objective of cross-national comparison is inextricably linked with global monitoring of whether the agendas reflected in international agreements (EFA, MDGs, SDGs, Paris Declaration on Aid Effectiveness, etc.) are being pursued at a national level. In a similar vein, governments voluntarily subject their educational systems to regional or cross-regional large-scale student assessments (PASEC, PIRLS, PISA, PISA-D, SACMEQ, TERCE, TIMSS, etc.) in order to position them internationally. Arguably, these international policy tools have a harmonisation, homogenisation and standardisation effect at the level of policy talk. Topics of great import in this regard, including what particular moments national policy actors become receptive to supranational influence, why they selectively adopt global education policy and how they adapt and modify them to suit their own agendas, are discussed elsewhere (see Steiner-Khamsi and Waldow 2012).

The rapport between politics and statistics is not a new field of inquiry. Desrosières (2000) examines official economic statistics in the liberal, welfare, Keynesian and neoliberal states. Cussó (2006) focuses on educational statistics and international organisations. Smyth (2008) studies the emergence of International Governmental Organizations and international classifications of education. Nóvoa and Yariv-Mashal (2003) link the current rebirth of educational comparatism to the rise of indicators in the intersection between technical and political spaces.

Bearing in mind the importance of context, we examined changes in education indicator research over time. The three periods selected for our study reflect important shifts in how and why education indicators were developed. In each case the following objectives apply: modernisation/nation-building; colonisation/development; and standardisation/globalisation. The first part of each pair (modernisation, colonisation, standardisation) reflects the policy objectives of educational statistics. The second (nation-building, development, globalisation) denotes the broader political and economic context in which the policy uses and abuses of cross-national comparison, educational statistics and indicator research were embedded.

2. Jullien’s scientific project: observe, compare, borrow

Jullien de Paris is considered one of the ‘pioneers’ of comparative education research (Sobe 2002, 141). Studying the genealogy of comparison as a ‘positive science’, that is, as an area that draws on specific questions, observations, measurements and tools of analysis, has developed into a fundamental topic of education research. But was Simon Fraser’s (1964, 115) suggestion that Jullien de Paris was the sole ‘father figure’ of comparative education correct? The 200-year anniversary of Jullien’s (1817) *Esquisse d’un ouvrage sur l’éducation comparée* marks the third time in which Jullien has been re-discovered as the founder of comparative policy studies in education. William W. Brickman, historian of education, comparativist and co-founder of the Comparative and International Education Society in 1956, reminds us that Jullien’s work was first found by chance in a bookstall along the Seine in Paris in the late-nineteenth century (Brickman 2010, 48). It was Ferenc Kemény, an influential Hungarian public intellectual, peace activist and promoter of international

education who came across Jullien's writings in 1885 (see Giraud 1975). Thereafter the publication date of Jullien's *Esquisse* became inscribed as the birthdate of the entire field of comparative education.

The second revitalisation of Jullien may be attributed to a faculty member at Teachers College, Columbia University, Isaac L. Kandel (1881–1965), a student of Sir Michael Sadler (1861–1943) who emigrated from the UK to the USA. For more than three decades (1915–1946) Kandel was a faculty member at Teachers College and for almost two decades (1946–1962) also served as consultant for UNESCO. In the meanwhile, Kandel (1942) effectively created a history for the field, which he believed had a promising future. Kandel's widely-cited publications led to the current, third revitalisation on the horizon for 2017, a time when cross-national comparisons and international large-scale student assessment are receiving unprecedented interest, inspiring a desire to delve more deeply into the history of indicator research, statistics and comparative policy studies in education.

Jullien de Paris was hardly the first to compare educational systems for purposes of drawing lessons, emulation and policy borrowing. In fact, the 'traveling observer' (Sobe 2002, 147) existed long before comparative education was established as a science. Gail P. Kelly (1992) refers to the early days of comparative education as the time 'when gentlemen traveled extensively and wrote about differences between nations' (14). In addition to these traveling observers there were also a few locally situated analysts who systematically compared their own systems with those abroad. In 1808, French Professor César-Auguste Basset urged his university to dispatch a researcher to 'foreign countries' to 'observe, compare, and present the facts' concerning their educational systems and methods of instruction (Brickman 2010, 47).

In Germany, Friedrich August Hecht has been credited with presenting the first comparative study 20 years prior to the publication of Jullien's *Esquisse* (Lenhart 2015). Hecht's *De re scholastica Anglica cum Germanica comparata* (English and German school education compared) was published in 1795 in Latin, which may help explain why, as with Jullien, his book was only discovered in the twentieth century once comparative education became established as an academic field requiring its own history and genealogy. However, there are two features of Jullien's work that set him apart from other early comparativists: in addition to Basset's above-mentioned dictum (observe, compare and present the facts), Jullien developed the methodology of *standardised* cross-national comparison and actively promoted policy borrowing. Thus, Jullien modified Basset's formula by encouraging researchers to borrow. From a critical perspective, one may assert that he developed indicators to replace the observer and researcher with an anonymous authoritative voice presenting the 'facts' or data for recommendations on what should be transferred from one national context to another.

Jullien had, as Gautherin (1993) poignantly describes, a 'mystical' belief in the power of public education to 're-generate' people by emancipating them from the feudal monarchical structures still dominating the political scene in France and Europe. In his early career Jullien was deeply impressed by the Swiss educators Johann Heinrich Pestalozzi and Philipp Emanuel von Fellenberg. Pestalozzi ran three institutes in Switzerland, where he implemented his child-centred pedagogy in the German speaking towns of Burgdorf and Münchenbuchsee, as well as in Yverdon, located in the French speaking part of Switzerland. Jullien moved with his family to Yverdon and lived in Pestalozzi's institute for three months. After this brief stay, he returned several times to Yverdon between 1811 and 1812 to more

fully study Pestalozzi's child-centred pedagogy. Partly because of Pestalozzi's positive influence on his mythical belief in public education as the means to bring happiness, mitigate social class differences and foster political participation among citizens, Switzerland became an important laboratory for Jullien's comparative studies. He also selected Switzerland because its decentralised governance system resulted in 22 Cantonal educational systems, roughly reflecting the wide variety found in the rest of Europe. As Jullien wrote:

Because of the great variety of climates, languages, religions, political organizations and governments in the twenty-two cantons of the Helvetic Confederation, an infinite variety of educational establishments and systems, reproducing every possible known form is to be found there. (Jullien de Paris 1817, 9)

Over the course of his travels, Jullien compiled a set of indicators to be used for comparison between different countries. He systematised and presented them as 'questions' in an Annex to the *Esquisse* (Jullien de Paris 1817). What strikes the reader even today is how much these indicators anticipate the vast majority of those currently used for international comparison. What has changed, however, are the definitions of an ideal educational system. The covert normative assumptions underlying the selection of suitable indicators account for changes over the three time periods under study in this article.

2.1. Jullien's visions of pedagogy

In developing his methodology, Jullien made two important assumptions. First, unlike some of his contemporaries, Jullien did not believe in 'the national character' of educational systems, and explored other factors to account for differences between systems. In fact, his studies in Switzerland – one nation with many different educational systems – demonstrated the limitations of the national-character theory prevalent at the time. Fraser (1964) contrasts Jullien's approach to interpreting differences between educational systems with that of Berchtold (1759–1809), who believed that educational systems reflect typical 'French', 'English' or other cultural features of their respective nation-states. Second, Jullien focused on formal, public schooling understood as a series of stages (primary, secondary, higher education) where different types of schools cater to the individual's evolving needs. As a result of these assumptions, the practice of comparing educational systems today begins with describing the organisations that impart education at the various levels, the populations they serve and the resources and strategies they employ.

The structure and organisation of the questionnaires laid out in *Esquisse* (Jullien 1817) reflect his vision of an ideal educational system that would help create the new citizen for a post-feudal, modern nation-state. Here, Jullien presented six 'series of questions' that contain the seed of contemporary indicators. Series 1 to 3 focused, respectively, on primary, secondary and higher education – a triad that mirrors the successive stages of childhood, adolescence and youth. Series 4 to 6 would have covered teacher education, the education of females and the regulatory framework, however the *Esquisse* presents only the first two series. Of the remaining four we only know their format and intent. Each of the first five series, those that deal with a subsector of the education system, follows the same blueprint: a triad of triads.

The first triad, the most 'quantitative', covers schools, teaching staff and students. Jullien prefigures indicators still used today; for instance, he asks 'What is the proportion of the total number of these students to the population of the township or division?' This foreshadowed

today's gross enrolment ratio (GER¹), which limits the denominator to school age children, instead of the whole population. Since Jullien also inquired about ages of school entry and exit (questions 34 and 40, first and second series), he might have taken this further step, had he implemented his project. But Jullien is also sensibly cautious. For example, after proposing a 'pupil-teacher ratio' (PTR), he adds: 'What is the minimum, the maximum?' (question 33, second series), because average PTRs disguise inequities (Kadzamira and Rose 2003), one of his recurring concerns.

The second triad, influenced by his observations in innovative European schools, focuses on curriculum and instruction. Its structure stems from Pestalozzi's tripartite conceptions of man and happiness, which constitutes the purpose of education: (1) physical and gymnastic education focuses on the body's health; (2) moral and religious education on the heart's morality and virtue; and (3) intellectual education on the mind's instruction.

The third triad focuses on each level's (e.g., primary) relationship with its context: the ties of public education with private or home-based education; its connections with other levels of public education; and its historical context. Here Jullien relaxes his own rules, namely the focus on public or formal education, as well as its compartmentalisation into levels.

2.2. The political agenda of compulsory education and nation-building

Arguably, education indicators tend to measure issues that, in a given policy context, are considered new, relevant or contested. Understanding the political agenda of a comparative study is therefore key to explaining why certain indicators are selected at the expense of others. Unsurprisingly, Jullien's indicators reflect his pedagogical mission, tailored after Pestalozzi's writings, as well as his belief in the modern nation-state where formal education is supposed to create social cohesion and mobility. Therefore the latter set of indicators corresponds with the urge of government actors at the time to establish [modern] schooling as a means of national-building. In effect, he combined pedagogy with politics in that his formula helped advance a very specific pedagogical vision in a post-feudal era where the political value of compulsory education for ensuring literacy among the masses was deemed a prerequisite for political participation and the spread of democratic values.

France's nation-building process, partly fuelled by evolving official statistics, surely inspired Jullien's political approach:

The most visible manifestation of this process of homogenising and codifying many aspects of human existence was the unification of the national territory, since many things and rules that were then redefined and generalised had previously been specified at a local or provincial level. This complex, expensive, and often tedious work was termed *adunation* by Sieyes, one of the principal actors involved: by this he meant the deliberate unification of systems of reference. ... The general principle was to make a tabula rasa of a society that had previously been characterized by its privileges, by fiscal arrangements that differed according to the province, and by local superstition. ... This process was launched by a monumental statistical survey. (Desrosières and Naish 1998, 32–33)

The modernising intention behind Jullien's plan is evident in several ways. First, it monitors the spread of innovative pedagogies and measures the prevalence of modern and traditional approaches. Second, it encourages exploration and experimentation to devise new methodologies for instruction, identifying what we would call 'best practices'. Third, it proposes a holistic view of human development, following Pestalozzi, and a systemic approach

to education. Finally, meliorism informs the entire project, within the Enlightenment's philosophical framework.

During his travels, Jullien observed innovative pedagogies in action, and encouraged adoption of ideas by Bell and Lancaster, Pestalozzi, as well as Gaultier (see Jullien de Paris 1817; first series, questions 96–98). Meanwhile, other questions in Jullien's exploratory/experimental approach are not data collection tools but interrogations on effective, ethical pedagogy: 'How does one succeed in directing and supervising, in an indirect way, without disturbing the freedom of children, their exercises and their games?' (Jullien de Paris 1817, question 54, first series).

Though featured in the physical education section, the question contains clear moral and intellectual implications illustrating Jullien's holistic approach. After inquiring about the use of rewards for good behaviour or competition to foster progress in study, Jullien proposes forsaking both in favour of focusing on a child's 'internal satisfaction ... in the consciousness of his strength and progress as he exercises and develops his moral character and his intelligence' (questions 86, 88 and 90, Fraser 1964). One may be tempted to propose that this oddly modern idea was popularised many years later and re-emerged most recently in the 1970s under the auspices of *intrinsic motivation* (Deci and Ryan 1975).

Jullien's underlying systemic thinking cuts across sections, exploring links between wealth, parenting practices and formal schooling. At one point, Jullien considers the relationship between social class, use of nursemaids and maternal influence on children's education. He first asks whether 'the wealthiest citizens commit their children to nursemaids' (question 43, physical education, first series); then if 'mothers exercise a considerable influence on the primary education of their children' (question 78, moral education); and, finally, if there are differences 'with regard to this influence of mothers, in the poor and rich classes' (question 79, first series). These types of questions reflect his mythical belief in the act of education, whether it takes place at home or in school, as a panacea for social inequality and feudal structures.

Another example of Jullien's systemic thinking is manifested in his meliorist approach. He was not only interested in describing and comparing systems – he wanted to improve them. For example, he links teachers' sociodemographic characteristics, experience and performance with students' learning outcomes. After inquiring about teachers' performance, age and experience (question 19, first series), Jullien wonders whether it is 'the younger ones or the older ones who are most successful and who form the best students?' (question 20, first series). In this instance, he ponders two mutually exclusive hypotheses: is good teaching explained by teachers' experience or by pedagogic innovations to which younger teachers may be more exposed and receptive? Jullien unequivocally comes down on the side of the latter by advocating for the innovative, student-centred pedagogical practices he expects to find more prevalent among younger teachers.

Jullien's ideological roots in the Enlightenment are evident throughout, as when he demands respect for 'the freedom of children' (question 54, first series). Equality is a recurring concern, expressed mostly in terms of the gap between the rich and the poor (questions 79, 112–114, first series). As for solidarity, Jullien asks whether adolescents in schools are 'stirred up against certain nations' or instilled with 'a universal kindness for men and even animals' (question 71, second series). Jullien's indicators were meant to determine what kind of public education succeeds in making the individual happy and the modern nation-state equitable.

After 1817 Jullien moved on to other professional interests and never published the remaining series of questions, abandoning his project on comparative education (Fraser 1964). Consequently, he did not leave behind specific indicators. But Jullien's project clearly entailed modern, equitable, holistic education as the foundation for building the French nation. France outperformed Switzerland in higher and scientific instruction, but the fact that French instruction, 'badly guided and badly divided, has not yet reached the lower classes ... has produced in part the faults, crimes and misfortunes whose reaction has spread throughout Europe' (Fraser 1964, 47). Jullien's example for France is Swiss 'elementary and communal education, the essential foundation of all education' (48). Therefore he cast a wide net with his questions, encompassing all domains of education and learning, including attitudes and values, and encouraging the 'lower classes' to abandon superstition and embrace a more modern religiosity.

3. Monroe's attempts to understand development stages

Although the objects of study have changed over the past two centuries, the method of cross-national comparison as a policy tool remains constant. For example, for seven years in the 1930s the International Examination Inquiry gathered eminent scholars from the USA, Scotland, England, France, Germany, Switzerland, Norway, Sweden and Finland to conduct comparative research on effective entrance examinations for secondary schools (Lawn 2008). One of the purposes was to learn from educational systems that had implemented policies tackling specific problems, such as the expansion of secondary schooling in Europe and the USA. The purpose, as conceived at the time, was to learn from educational systems that were more 'developed'. The strong belief in development stages is best illustrated by Teachers College faculty who participated in cross-national studies comparing the USA and Europe and advised governments who commissioned them to review the education sector and provide recommendations for reform. Among them was the most internationally renowned faculty member of Teachers College at the time, Director of the International Institute, Paul Monroe.

3.1. Monroe's survey: understanding local needs

Paul Monroe's 'surveys in foreign lands', conducted in the early-twentieth century, showcased changes in the use of indicator research and education statistics, for instance in the Philippines while they were under American tutelage (see Monroe, Páez, and Duggan 1925). This work took place after the creation of the permanent Bureau of the US Census in 1902, in an era of official statistics characterised by 'renewed relations between the administration and the academic community' (Desrosières and Naish 1998, 199)

As a faculty member in comparative education and history of education at Teachers College, Monroe taught three groups of students (see Bu 1997): American missionaries who believed in a universal method of education based on the Bible; secular students aiming to work for the colonial administration of the USA, other colonial empires or as experts for foreign governments; and international students who intended to return to their countries as education leaders who would adapt lessons from the USA and Europe for the schooling of their citizenry.

In addition to teaching and research at Teachers College, Monroe advised governments based on ‘surveys’, that is, empirical investigations of their educational systems. The main purpose of Monroe’s survey was to identify local needs. Nowadays, this would be considered an education sector review because he used a set of indicators to describe and analyse a given educational system’s status quo. Monroe’s commitment to understanding local needs was aligned with the concept of Adapted Education, used in the USA for the schooling of Native Americans as well as African Americans in the South, which was then exported by US missionaries and academics to other parts of the world. The British colonial empire also mandated Adapted Education for its colonies during the same period (see Steiner-Khamsi and Quist 2000). Understanding local needs sharply contrasted with the traditional missionary practice of disregarding local circumstances to teach a curriculum proselytising the same universal Christian beliefs. Using indicators and statistics to describe and compare educational systems was labelled the ‘positive method’ or, during Monroe’s time, the ‘scientific method’ in education.

3.2. The belief in development stages and contextual comparison in twentieth century comparative education

The longstanding practice of comparative researchers warning, in theory, against de-contextualisation, while at the same time advancing, in practice, educational transfer across contexts, raises specific questions: How have comparativists dealt with the discrepancy between theory and practice in comparative education? In particular, how have they justified educational transfer? Early-twentieth century comparativists such as Monroe (USA), Kandel, as well as his mentor, Sadler, in the UK, used Civilisation Theory to construct comparability. Individuals, groups, or nations perceived to be at the same stage of ‘civilisation’ were deemed comparable. Most commonly, researchers identified the particular development stage for each of the ‘races’, nations and educational systems. African Americans, native Americans and Africans (and later all colonised people in the British colonial empire), were regarded as being at the same underdeveloped stage of civilisation. Since these groups were thus comparable, transfer of educational models from one continent (North America) to another (Africa) was viewed as methodologically and morally permissible. For example Jesse Thomas Jones, chairman of the Phelps-Stokes African Education Commission, justified the transfer of the American Hampton-Tuskegee model to the African continent as follows:

[T]hough village conditions in Africa differ in many respects from those in America, where these activities [of Hampton and Tuskegee] had great influence on the improvement of rural life, the resemblances are sufficiently numerous and real to warrant the belief that the plans above described may be adapted to colonial conditions in Africa. (Jones 1922, 141)

A good illustration of the (development) stage model pursued in comparative education in the first half of the twentieth century is manifested in the work of Kandel. Kandel advocated for a contextual study of educational systems called the historiographic method in comparative education. The *Education Yearbook*, which he edited between 1924 and 1944, consisted of a compilation of historiographic ‘one-country studies’. The yearbook was developed with the intention of identifying the ‘national character’ presumably underlying each educational system (see Pollack 1993).

Following the stage model, Kandel only compared educational systems that he believed were at the same level. This, to take one example, justified the transatlantic transfer of

educational models that dealt with nation-state building and compulsory schooling between the newly formed nation-states in Europe and the USA, was justifiable. This is not to exclude, however, lesson drawing from 'less developed educational systems'. In 1944, Kandel wrote:

The development of education appears in most countries to have followed the same rhythm. The nineteenth century opened with a movement to establish systems of universal and compulsory education. (Kandel 1944)

According to Kandel and some of his contemporaries, the next stage in the development of educational systems following the establishment of universal and compulsory basic education, was the expansion of secondary schooling, followed by the implementation of vocational and technical education, followed by the establishment of a system for higher education.

In the first half of the twentieth century, scholars in North America and Europe discovered Africa, Latin America and East Asia as laboratories for social research and educational studies. They were fascinated by what European and North American educational systems could learn from less developed countries, and what the poor countries, in turn, could adopt from past experience in education and nation-building in their richer, more 'civilised' counterparts. For example, in 1944 Kandel contended that twentieth century reform in Jamaica could benefit from lessons learned in the UK a century earlier. With great compassion for the difficulty of implementing a fundamental educational reform in an island nation that was still part of the British empire, Kandel noted that the Jamaican educational system had 'to meet all these demands for education at one time, while other countries have met them slowly and piecemeal' (Kandel 1944 [reprinted 1999], part 1). He encouraged educational officials in Jamaica to learn from experiences of other educational systems, particularly England's:

From the point of view of the administration and organisation of education, Jamaica stands today [1944] in the same position as did England in the last years of the nineteenth century. (Kandel 1944)

Renowned scholars such as John Dewey were excited by the idea of turning back the clock on educational development to avoid the mistakes of 'educationally advanced' countries and transferring an improved version to those which were 'educationally new'. Unlike most scholars at the time, Dewey hesitated to transfer educational models from 'educationally advanced countries' to 'less civilised', 'less complex' and 'less developed countries'. However, Dewey's visit to the Mexican rural education programme in 1926 was a turning point. Upon his return, he 'revived his faith' in educational transfer, fascinated by the opportunities for 'educationally new' nations to begin 'starting afresh with the most enlightened theories and practices of the most educationally advanced countries' (cited in Goodenow 1990, 29).

More recently, processes of decolonisation in Africa reinvigorated the late-eighteenth century notion of *adunation*, which 'became a global template in the second half of the twentieth century'; this *methodological nationalism* gave African leaders a framework to use statistics in building their nation states, and later became the foundation of the MDGs (see Speich Chassé 2016, 211).

4. SDG indicators: the use of global norms for monitoring national progress

The third period, starting in the 1990s, is characterised by globalisation, defined as the increased transnational flow of people, production, finance, ideas, beliefs and technology, accompanied by processes of policy harmonisation, homogenisation or standardisation. In the global education landscape the establishment of international conventions, such as EFA, MDGs, SDGs, and the rise of international organisations (UNESCO, World Bank, OECD, etc.) that (selectively) translate and enforce the conventions by operationalising, measuring, monitoring and evaluating them, have had the greatest impact on which new indicators have been formulated. That said, it is important to bear in mind that the new set of indicators were merely added to the already existing set of commonly used indicators, formulated and used over the past 200 years by Jullien, Monroe and many others. Nóvoa and Yariv-Mashal (2003) see current educational comparatism as partly fuelled by indicators meant to promote reforms or legitimise policies, while they wonder if the future of the discipline will revolve around modes of governance or scientific concerns. Similarly, Cussó (2016, 56) aims ‘to bring to light the subtleties of balancing accuracy (methods) with institutional needs (policies)’, elucidating how data are produced and what purpose they serve.

International organisations such as UNESCO, the World Bank, OECD and so on function on the one hand as norm setters, and on the other as monitors and evaluators of national progress towards the norms or benchmarks that they had established with their member states or constituents. As with the two earlier periods, it is important to analyse the policy context – which we today call the *global* policy context – to understand why some indicators were newly developed or added to the already existing sets of indicators.

Our focus is neither a meta-analysis of the statistics nor their use in politics (Cussó 2016), but their intersection, making indicators for international commitments particularly enlightening. For example, EFA Goal 4 calls for a 50% improvement in adult literacy rates, while target 5.5 aims to ‘Reduce illiteracy rates’ by 50% (UNESCO 2000, 31). Allegedly, decision-makers in Dakar reversed the latter ‘negative’ target wording into the former ‘positive’ goal; however, the formulations are not mathematically equivalent, with the final one deemed ‘arithmetical nonsense’ (Guadalupe 2015). Remarkably, for countries with initial literacy rates below 50%, the ‘positive’ goal is *less* ambitious, but arguably more realistic, than the ‘negative’ target. For example, in 2001 Niger’s adult literacy rate was 14%.² For 2015, the ‘positive’ target would be 21%; the ‘negative’ target, 57%. The 2015 rate was 19%. Does this reflect politicians’ inability to understand the technocrats’ logic? Or the *reverse*?

Clearly, two phenomena have achieved unprecedented prominence since the 1990s: international commitments (MDGs, EFA) monitored with data produced by the UNESCO Institute for Statistics and analysed by the Global Monitoring Report; and the growth of international large-scale assessments (ILSAs). However, the two did not converge until the recent inclusion of learning outcomes in the SDGs. Founded in 1999, the UNESCO Institute for Statistics (UIS) has been, in comparison to OECD, the World Bank and other international knowledge banks, something of a latecomer – particularly given its relatively recent shift towards learning outcomes. The focus on global competitiveness of national education systems – as measured in students’ learning outcomes – is what drives the politics of education statistics in the new millennium. Desrosières described a recent shift in international statistics, regardless of sector, ‘from the harmonisation of outputs (governments classifying their data in accordance with international standards) to the harmonisation of

methods (governments producing internationally standardised statistics at source)' (Cussó 2016, 59). Is this happening in education? If so, do the creation of the UIS, the rise of ILSAs and the advent of the SDGs play a role here?

In the 1990s 'UNESCO's Division of Statistics continued to measure and compare the spread of mass education and literacy' to spur development while respecting cultural identity (Cussó 2006, 533). Meanwhile, organisations like THE OECD and the World Bank criticised UNESCO's approach while demanding data on learning achievement. For some, UNESCO's *development comparatism* stood in contrast to the OECD's *globalisation comparativism* (Cussó & D'Amico 2005). These debates led to the creation of the UIS in 1999 (Cussó 2006). The UIS has since reported on globalising processes, such as the international mobility of higher education students. In addition, over the years the UIS has increasingly focused on the efficiency of education systems (UIS 2003–12), which is consistent with *globalisation comparativism* (Cussó and D'Amico 2005). However, the UIS did not disseminate learning assessment data (UIS 2003–2012). Why not? After all, EFA goal 6 did call for, among other things, 'measurable learning outcomes ... especially in literacy, numeracy, and essential life skills'; however, 'the Millennium Development Goals, which have received greater attention over the past 15 years, were more narrowly focused on access' (Rose 2015, 487).

An example of the standardising trend in education is the growth of large-scale assessment, which happened simultaneously at the national, regional and cross-regional levels: countries shifted their focus from examinations to national large-scale assessments (NLSAs); three regional assessments emerged; and cross-regional assessments became widespread and periodical.

At the country level global actors fuelled a shift from the golden age of examinations (1963–1993) to NLSAs, which 'were by far the most common type' supported by the World Bank in 1998–2009 (Lieberman and Clarke 2011). Unlike examinations, NLSAs aim to provide information on system performance and associated factors (Clarke 2012), thus sharing purpose and methodologies with ILSAs. Arguably, implementing NLSAs prepares countries to join ILSAs, and vice versa.

Meanwhile, regional assessments, neither national nor global, defy classification: they are included in analyses of ILSAs by some authors, (Pizmony-Levy 2013), but not by others (Heyneman and Lee 2013).³ Some link them to national assessments, extolling their virtues: 'regional and national assessments can be advantageous when examining the quality of education and the curriculum that is actually delivered by schools and teachers. They tend to be more context-sensitive' (Kamens and Benavot 2011, 286).

International large-scale assessments started in the 1960s with the International Association for the Evaluation of Educational Achievement (IEA).⁴ Since the 1990s, new cross-regional assessments reach a larger, more diverse set of countries and adopted 'predictable (?) cycles' (Pizmony-Levy 2013, 58), enabling periodical updates on progress.⁵ This, coupled with OECD's entry into the field, prompted a shift 'from research to policy' (Wagemaker 2013, 11), towards 'audit and accountability' (Pizmony-Levy 2013, 54).

Nowadays, ILSAs are known for their controversial rankings by mean or average score, which are changing *how* educational indicators are produced. This deviates from the traditional rates and ratios as the basis for indicators on enrolment, promotion, literacy and so on.⁶, which distinguish between *haves* and *have nots* to guarantee access to education, if not necessarily learning, for everyone. By contrast, averages suit a focus on global competitiveness because they are sensitive to extreme values, on both ends of the distribution,

that is, both low- and high-performing students exert a pull on the mean value. PISA 2009 aimed ‘to identify highest-performing students’ (OECD 2010) by finding their common features, thus extending ‘best practices’ from countries to individual learners.

A major challenge for ILSAs is alignment with curriculum. There are two competing approaches: IEA analyses participating countries’ curricula and distils their commonalities into the test (Martin and Kelly 1996); whereas PISA assesses skills for solving ‘real-life’ problems, regardless of curricula (OECD 2010). Both assume that, across participating countries, either schooling or real life has become sufficiently homogeneous to justify a single measurement tool. Critics argue that a focus on core skills may lead countries to neglect more idiosyncratic areas, such as social sciences, not covered by ILSAs. However, there is no evidence of an *independent* effect of ILSAs on policy; instead, national actors exploit results according to their own goals (Teltemann and Klieme 2016). The question is: Could this still, eventually, and in patches, depending on national actors, lead to a feedback loop of cross-national standardisation of curricula? And what role will SDG 4.1.1⁷ play by emphasising learning outcomes and explicitly focusing on reading and mathematics?

In sum, the country-level shift from examinations to system-level assessments, the advent of regional assessments and the rise of periodical cross-regional assessments all contributed to making large-scale, standardised assessments a defining trait of the current global education landscape. In this process, averages joined rates and ratios among the most visible statistical devices for producing educational indicators. Finally, questions arise about ILSAs, and the SDGs, as possible influences on national curricula.

This rise of assessments gained even more momentum when testing met tracking. In July 2012, as the international community prepared for the end of the MDG and EFA commitments, UIS and the Brookings Institute launched the Learning Metrics Task Force. This global network effectively advocated a shift ‘from a focus on access to access plus learning’ (LMTF 2013, 1), culminating in the inclusion of learning outcomes in the SDGs.

Since 2014, the ongoing SDGs debate, increasingly technical and inevitably political, has shifted from goals to targets to indicators. The next step requires consensus on data sources. According to the UIS, this may be challenging, since ‘there is currently no framework to reconcile the differences between the various types of assessment to produce cross-nationally comparable data’ (UIS 2016, 57).

Conversely, former Director of the UNESCO Global Monitoring Report, Pauline Rose, warns against comparability at the expense of national needs:

An international curriculum based on a global system of international large-scale assessment could further result in a misalignment between the education system and a country’s needs in relation to its stage of economic development and labor market demands. (Rose 2015, 488)

This tension may be resolved through a gradual transition towards standardisation. According to the Inter-Agency Expert Group on SDG indicators, the sources for indicator 4.1.1 will be cross-regional, regional, national and citizen-led learning assessments. Although these are not mutually comparable, tracking will be based on individual studies while common scales are created over the next few years.⁸ This gradual transition from ‘individual studies’ to ‘common scales’ evokes the shift from harmonising outputs to harmonising methods (Cussó 2016).

5. Conclusion

The discursive power attached to numbers is not to be underestimated. In terms of Jullien's methodological contribution, one may invert the causal relation between indicators and comparability, which is commonly assumed in the comparative inquiry of national educational system. The use of indicators *makes* educational systems comparable regardless of how different they are. Put differently: educational systems are not comparable *per se*, but they are made comparable by using the same measurement. Today, turning features of a few educational systems into a global set of indicators that allow researchers to compare but also to monitor progress has become a cause for celebration for some, and concern for others. Advocates for evidence-based policy planning view indicators as an effective policy tool to advance political, social and economic accountabilities in educational development. Critics, however, decry the pace of homogenisation and standardisation that have resulted from global norm-setting (see Mundy et al. 2016).

It is important to bear in mind that 'statistics', meaning the science of the state (Foucault 1991, 96), became an important tool for the governmentalisation of the state ('governmentality') and is considered a project that, according to Foucault, started in European countries in the sixteenth century and expanded during the time of nation-building in the eighteenth and nineteenth centuries. The evidence suggests that gathering data in a nation and comparing data across nations was fashionable during Jullien's time. Educational statistics was an important yet understudied feature of modern schooling that the emerging nation-states had created. The project of the modern nation-state to amass data on its citizens is a precursor to today's knowledge banks of international organisations, such as OECD, the World Bank, UNESCO and other international organisations. Today, the international knowledge and databanks not only collect statistical information on their member states, but also include portfolios of 'best practices', often presented as international standards, which they readily disseminate, in some cases with funding, across the globe.

What statistics did before at the national level, they now do at supra-national level: statistics enable institutions to regulate and sanction. Today, the means of regulating and sanctioning have changed from a top-down approach to one of constant self-monitoring and self-evaluation, exacerbating the reliance on data for measuring progress. As a result of 'soft governance', institutions at all levels comply with international standards and agreements for fear of being left out or left behind. There is a vast body of emerging studies on 'soft governance' and 'governing by numbers' that has drawn the attention of a wide array of scholars, including those who examine the new modes of governance in the post- or neo-bureaucratic state (Maroy and Voisin 2014; Ozga 2016) and those who describe the shift from bureaucracies, accountable to an 'external sovereign with roots in authority and power', such as the state, the church, associations and so on, to so-called 'hyper-organisations' accountable to a set of agreed-upon standards (Bromley and Meyer 2015, 11).

What made Jullien stand out *vis-à-vis* other diligent empirical educational researchers and comparative social scientists of his time was precisely his usage of data for informing policy decisions. As Jacqueline Gautherin (1993, 2) has poignantly remarked, Jullien saw education, or more precisely schooling, as a vehicle for individual happiness and for a more egalitarian society. As a result, Jullien actively promoted policy borrowing across European countries. For him, the use of indicators for lesson-drawing and policy borrowing has been, from the onset, inextricably linked to the 'scientific method of comparison'.

An innovation does not on its own guarantee adoption and recognition by others. One would have to immerse oneself in the history of ideas during the eighteenth century to examine in greater depth the relation between Jullien and other early comparative thinkers to determine how and why Jullien's contributions have endured into the present. For example, the most well known early European comparative education researcher, director of public instruction in France, Victor Cousin (1792–1867) was not only able to carry out educational reform, but helped spin a policy dimension to the comparative inquiry of educational systems (see Brewer 1971). Cousin followed in the footsteps of Jullien de Paris in actively promoting policy borrowing across national boundaries and contexts.

Jullien's project, never fully operationalised, paradoxically preserved its complexity and ambition. His questionnaire, without actual indicators, cast a wide net that reflected his holistic, modernising educational approach aimed at building an egalitarian, enlightened France. Since then, Jullien's legacy has endured and evolved: enrolment rates and other indicators of access were treated as markers of development throughout the twentieth century; nowadays, increasingly standardised cross-national measures, such as average scores in ILSAs, are regarded by some as predictors of future global competitiveness and economic growth.

Historically, educational comparisons have conceptualised world regions in different ways. Jullien focused only on his own continent, ignoring most of the world, as he encouraged European nation states to learn from one another. Monroe and his contemporaries broadened this view, but cast two groups of countries in different roles: less civilized countries should emulate more civilized ones, without repeating their mistakes; conversely, less civilized countries provided a laboratory where more civilized countries could learn from this experimentation. Clearly, there was a colonial map underlying early-twentieth century comparative education research. Finally, in the era of the SDGs, global norms and benchmarks are not associated, in theory, with particular regions, but with devices such as ILSAs that showcase models worth emulating. That said, the most influential ILSAs originated in the developed world and usually single out industrialised models.

The SDGs, and particularly indicator 4.1.1 on student achievement, have sparked an ongoing debate about which sources of data should be used to monitor learning outcomes at the global level: national, regional, cross-regional assessments or some combination thereof? This discussion involves technical issues pertaining to data quality, such as coverage, validity and reliability. But it also touches upon cultural and political issues, notably the alignment (or lack thereof) between assessment, at any level, and curriculum, which up until now has been designed at the national level (or at lower levels). Thus, the need for curricular alignment at the country level creates an inevitable tension with the global actors' demand for comparability. Once again, as in Jullien's time, comparability must be *made*, certainly not from scratch, but in accordance with current beliefs about education and bearing in mind that by somehow making systems comparable, this global process may influence the mechanisms by which education systems appraise themselves and, as a result, the types of education systems they will strive to become.

Notes

1. The GER is calculated by dividing the number of enrolled children by the number of school-age children in the population. <http://uis.unesco.org/sites/default/files/documents/education-indicators-technical-guidelines-en.pdf>
2. <http://uis.unesco.org/country/ne>
3. PASEC started in 1991; SACMEQ in 1995; LLECE in 1997.
4. The IEA studies never cast themselves as 'regional'. However, historically, most participating countries were industrialised, a group treated as a 'region' by most international organisations.
5. The IEA launched TIMSS and PIRLS in 1994 and 2001, respectively; the OECD launched PISA in 2000.
6. <http://uis.unesco.org/sites/default/files/documents/education-indicators-technical-guidelines-en.pdf>
7. SDG Indicator 4.1.1: Proportion of children and young people: (1) in Grade 2 or 3; (2) at the end of primary education; and (3) at the end of lower-secondary education achieving at least a minimum proficiency level in (1) reading and (2) mathematics, by sex.
8. <http://unstats.un.org/sdgs/files/metadata-compilation/Metadata-Goal-4.pdf>.

Disclosure statement

The authors have no conflicts of interest to declare. The information of this document expresses the authors' personal views and opinions and does not necessarily represent UNICEF's position.

Funding

This research received no funding from UNICEF or any other agency.

ORCID

Manuel Cardoso  <http://orcid.org/0000-0002-2706-9839>

References

- Brewer, Walter V. 1971. *Victor Cousin as a Comparative Educator*. New York: Teachers College, Columbia University, Teachers College Press.
- Brickman, William W. 2010. "Comparative Education in the Nineteenth Century [Reprinted]." *European Education* 42 (2): 46–56.
- Bromley, Patricia, and John W. Meyer. 2015. *Hyper-Organization: Global Organizational Expansion*. Oxford University Press.
- Bu, Liping. 1997. "International Activism and Comparative Education: Pioneering Efforts of the International Institute of Teachers College, Columbia University." *Comparative Education Review* 41 (4): 413–434.
- Clarke, Marguerite. 2012. What matters most for student assessment systems: A framework paper. SABER-Student Assessment Working Paper, 1. Washington, DC: The World Bank.
- Cussó, Roser. 2006. "Restructuring UNESCO's Statistical Services – The "Sad Story" of UNESCO's Education Statistics: 4 Years Later." *International Journal of Educational Development* 26 (5): 532–544.
- Cussó, Roser. 2016. "From Statistics to International Quantification: A Dialogue with Alain Desrosières." In *The Social Sciences of Quantification*, edited by Isabelle Bruno, Florence Jany-Catrice, and Beatrice Touchelay, 55–65. Cham, Switzerland: Springer International Publishing.
- Cussó, Roser, and Sabrina D'Amico. 2005. "From Development Comparatism to Globalization Comparativism: Towards More Normative International Education Statistics." *Comparative Education* 41 (2): 199–216.

- Deci, Edward L., and Richard M. Ryan. 1975. *Intrinsic Motivation*. New York: John Wiley & Sons Inc.
- Desrosières, Alain. 2000. "L'État, le marché et les statistiques. Cinq façons d'agir sur l'économie." [The State, the Market, and Statistics. Five Ways of Acting on the Economy.] *Courrier des statistiques* 95–96: 3–10.
- Desrosières, Alain, and Camille Naish. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, Massachusetts: Harvard University Press.
- Foucault, Michel. 1991. "Governmentality" [original 1978, lecture given at the Collège de France]. Reprinted in *The Foucault Effect*. In *Studies in Governmentality*, edited by Graham Burchell, Colin Gordon and Peter Miller, 87–104. Chicago, IL: University of Chicago Press.
- Fraser, Steward. 1964. *Jullien's Plan for Comparative Education, 1816–1817*. *Comparative Policy Studies*. New York, NY: Teachers College Press.
- Gautherin, Jacqueline. 1993. "Marc-Antoine Jullien ('Jullien De Paris'), 1775–1848." *PROSPECTS-UNESCO* 23 (3–4): 757–773.
- Giraud, par Jean. 1975. "Marc-Antoine Jullien de Paris (1775-1848)." *Paedagogica Historica* 15 (2): 379–405.
- Goodenow, Ronald K. 1990. "The Progressive Educator and the Third World: A First Look at John Dewey." *History of Education* 19 (1): 23–40.
- Guadalupe, César. 2015. *Contar para que cuente: una introducción general a los sistemas de información educativa*. Lima: Universidad del Pacífico.
- Heyneman, Stephen P., and Bommi Lee. 2013. "The Impact of International Studies of Academic Achievement on Policy and Research." In *Handbook of International Large Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, edited by Leslie Rutkowski, Mathias von Davier and David Rutkowski, 37–74. Boca Raton, FL: CRC.
- Jones, Thomas J. 1922. *Education in Africa: A Study of West, South, and Equatorial Africa by the African Education Commission, under the Auspices of the Phelps-Stokes Fund and Foreign Mission Societies of North America and Europe*. New York: Phelps-Stokes Fund.
- Jullien de Paris, Marc-Antoine. 1817. *Esquisse d'un ouvrage sur l'éducation comparée*. Translated and commented version reprinted in Steward Fraser (1964). *Jullien's Plan for Comparative Education, 1816–1817*. *Comparative Policy Studies*. New York: Teachers College Press.
- Kadzamira, Esme, and Pauline Rose. 2003. "Can Free Primary Education Meet the Needs of the Poor?: Evidence from Malawi." *International Journal of Educational Development* 23 (5): 501–516.
- Kamens, David H., and Aaron Benavot. 2011. "National, Regional and International Learning Assessments: Trends among Developing Countries, 1960–2009." *Globalisation, Societies and Education* 9 (2): 285–300.
- Kandel, Isaac L. 1942. "International Cooperation in Education: An Early Nineteenth Century Aspiration." *In the Educational Forum* 7 (1): 23–29.
- Kandel, Isaac L. 1944. "The 1944 Kandel Report on Education." Reprinted in *The Jamaica Gleaner*, Kingston, Jamaica, October 19 and 20, 1999.
- Kelly, Gail P. 1992. "Debates and Trends in Comparative Education." In *Emergent Issues in Education: Comparative Perspectives*, edited by Robert F. Arnove, Philip G. Altbach and Gail P. Kelly, 13–22. Albany: SUNY Press.
- Lawn, Martin, ed. 2008. *An Atlantic Crossing? The Work of the International Examination Inquiry, Its Researchers, Methods and Influence*. Oxford, GB: Symposium Books Ltd.
- Lenhart, Volker, ed. 2015. *Die erste Schrift zur Vergleichenden Erziehungswissenschaft/The First Treatise in Comparative Education. Fridericus Augustus Hechius: De re scholastica Anglica cum Germanica Comparata (1795 – 1798)*. Frankfurt: M: Lang.
- Liberman, Julia, and Marguerite Clarke. 2011. *Review of World Bank Support for Student Assessment Activities in Client Countries*. Washington: The World Bank.
- L.M.T.F. 2013. *Toward Universal Learning: Recommendations from the Learning Metrics Task Force*. Montreal and Washington DC: UNESCO Institute of Statistics and Center for Universal Education at Brookings.
- Maroy, Christian, and Annelise Voisin. 2014. "Une typologie des politiques d'accountability en éducation: l'incidence de l'instrumentation et des théories de la régulation." [A Typology of

- Accountability Policies in Education: The Impact of Instrumentation and Theories of Regulation.] *Education Comparée* 11: 31–58.
- Martin, Michael O., and Dana L. Kelly. 1996. *Third International Mathematics and Science Study. Technical Report. Volume I: Design and Development*. Chestnut Hill: Boston College.
- Monroe, Paul, José Páez, and Stephen P. Duggan. 1925. *A Survey of the Educational System of the Philippine Islands by the Board of Educational Survey*. Manila: Bureau of Printing.
- Mundy, Karen, Andy Green, Bob Lingard, and Antoni Verger, eds. 2016. *Handbook of Global Education Policy*. Somerset, GB: John Wiley & Sons.
- Nóvoa, Antonio, and Tali Yariv-Mashal. 2003. “Comparative Research in Education: A Model of Governance or a Historical Journey?” *Comparative Education* 39 (4): 423–438.
- OECD. 2010. *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science*. I vols. Paris, France: OECD.
- Ozga, Jenny. 2016. “Trust in Numbers? Digital Education Governance and the Inspection Process.” *European Educational Research Journal* 15 (1): 69–81.
- Pizmony-Levy, Oren. 2013. *Testing for All: The Emergence and Development of International Assessment of Student Achievement, 1958–2012*. PhD diss., Indiana University.
- Pollack, Erwin. 1993. “Isaac Leon Kandel.” *PROSPECTS-UNESCO* 23: 775–787.
- Rose, Pauline. 2015. “Is a Global System of International Large-Scale Assessments Necessary for Tracking Progress of a Post-2015 Learning Target?” *Compare: A Journal of Comparative and International Education* 45 (3): 486–490.
- Smyth, John A. 2008. “The Origins of the International Standard Classification of Education.” *Peabody Journal of Education* 83 (1): 5–40.
- Sobe, Noah W. 2002. “Travel, Social Science and the Making of Nations in Early 19th Century Comparative Education.” In *Internationalisation: Comparing Educational Systems and Semantics*, edited by Marcelo Caruso, 141–166. Berlin/Bern/New York: Frankfurt am Main.
- Speich Chassé, Daniel. 2016. “The Roots of the Millennium Development Goals: A Framework for Studying the History of Global Statistics.” *Historical Social Research* 41 (2): 218–237.
- Steiner-Khamisi, Gita, and Hubert O. Quist. 2000. “The Politics of Educational Borrowing: Reopening the Case of Achimota in British Ghana.” *Comparative Education Review* 44 (3): 272–299.
- Steiner-Khamisi, Gita, and Florian Waldow, eds. 2012. *World Yearbook of Education 2012: Policy Borrowing and Lending in Education*. New York: Routledge.
- Teltemann, Janna, and Eckhard Klieme. 2016. “The Impact of International Testing Projects on Policy and Practice.” In *Handbook of Human and Social Conditions in Assessment*, edited by Gavin Brown and Lois Harris, 369–386. New York: Routledge.
- UIS. 2016. *Sustainable Development Data Digest: Laying the Foundation to Measure Sustainable Development Goal 4*. Montreal: UNESCO-UIS.
- UNESCO. 2000. *The Dakar Framework for Action: Education for All: Meeting Our Collective Commitments: Including Six Regional Frameworks for Action*. Paris: UNESCO.
- Wagemaker, Hans. 2013. “International Large-Scale Assessments: From Research to Policy.” In *Handbook of International Large Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, edited by Leslie Rutkowski, Mathias von Davier and David Rutkowski, 11–36. Boca Raton, FL: CRC.