Chapter 16

The Politics of Metrics in Education: A Contribution to the History of the Present

Romuald Normand

16.1 Introduction

Policy instruments are linked to the development of new modes of governance. They provide cognitive and normative frameworks for policy-makers to advocate changes, to implement new programmes, and to create new types of public interventions (Lascoumes and Le Galès 2007). They contribute to the transformation of the State through the invention of new tools and devices, particularly metrics, which give legitimation to political aims, values, and ideologies. It corresponds to New Public Management which pretends reinventing tools of government and overcoming bureaucracy sometimes by reusing recipes from the past (Hood 1986). The instruments participate also in a kind of depoliticization and re-politicization of decision-making whereas policy-makers face many contestations and oppositions from different interest groups. As Michel Foucault demonstrated, these technical procedures of power and instrumentation are central to the art of governing and the development of a rationalizing State (Foucault 1977). Governmentality is not only based on measuring devices but also on intellectual and scientific technics, ways of thinking, epistemologies which become operational through metrics (Miller and Rose 2008). New relationships are established between science, expertise and politics that impact on the ownership, selectivity and choices of tools and instruments.

As a policy instrument, metrics are invading the area of education. Tests, indicators, and benchmarks are the panoply of New Public Management while teachers and students are daily exposed to accountability technologies. A political rationality is dissimulated behind this frenzy of calculation and comparison across time and space since the creation of modern education systems (Lawn 2013a, b). This posi-

R. Normand (✉)

Faculty of Social Sciences, University of Strasbourg, Strasbourg, France

e-mail: rnormand@unistra.fr

© The Author(s) 2020

G. Fan, T. S. Popkewitz (eds.), *Handbook of Education Policy Studies*,
https://doi.org/10.1007/978-981-13-8347-2_16

345

tivity, to resume Michel Foucault's words (2002), is clearly due to the effort of rationalization achieved through *mathesis* and *taxinomia* penetrating the spheres of the State administration, at least when statistics were used for governing population in education and health as in other economic and social areas (Porter 1996). Today, discourses of truth seize numbers, sometimes in caricatural ways subverting them to ideological and political aims (Berliner and Biddle 1995). Numbers speak by themselves, to reveal a truth while they are exhibited and imposed without any contestation by expertise and centres of calculation (Lawn 2013a, b). As the PISA survey, this methodological rationale is as a guarantee used by policy-makers to reject any criticism beyond the small circle of influential experts (Grek 2009, 2013). However, there is no naturality in the politics of metrics, no discourses and truths that could be established forever in education. Sociology and anthropology have contested these postulates and statements for long showing the relativity of data, their contextual embeddedness, and the cultural differences they tend to erase. But government by numbers appears legitimate for itself and politics of metrics become a new science of government at international and European level (Lawn and Normand 2014).

This chapter is resolutely critical but it is not contesting data or participating in the quarrel of methodologies, nor discussing the rightness of metrological arguments and justifications. It provides a singular argument on the politics of metrics in education through a social and political epistemology as an archaeology or history of the present (Foucault 2012; Popkewitz 1997, 2013). Historicizing the politics of metrics, its ideas and instruments, analytical models and theoretical frameworks, impacts on human beings, is adopting a critical perspective and considering education policies guided by the process of Reason (Popkewitz 2012). The aim is also to show the force of representations and power relationships leading to the adoption of such a scientific theory, an instrument, a methodology, a statistical artefact or metrological convention for States, local authorities and International Organizations.

We have chosen to structure this history of the present with examples characterizing the relationships between metrics, knowledge and politics since the beginning of compulsory schooling. We do not seek to establish continuities but to analyse some historical moments with their internal coherence, built on epistemological and discursive configurations underpinned by a certain kind of instrumentation impacting on politics. We highlight some concepts, theories, objects shaping rules of objectivity and scrutinize discourses of political truth. We look at changes within the areas of science and expertise and the regimes of normativity they introduce in the field of education policies.

16.2 The Politics of Classifications

From Michel Foucault (2002), we know the role played by natural history in the classification of human beings and things which structure the scientific language and announce the venue of the comparative chart. Classifications and comparisons are two elementary acts of any scientific approach and the former is the foundation

of metrics. The continuous, ordered and universal chart, of all the possible differences, is the ideal of *Taxinomia*. In addition, *Mathesis* defines a perspective for understanding the world from single laws stemming from the mathematical method. Since, a part of educational sciences has taken natural sciences as a model in attempting to reach the “perfection” of correspondence between truth and facts (Popkewitz 2012). Classifications are part of a politics which, by force or negotiation, facilitates the convergence between heterogeneous systems and conceptions. And even if they raise some ethical concerns, these modes of classification are finally made invisible and embedded in social and political life routines. Historically, it was the case of the classification of feeble-minded people as a policy instrument.

16.2.1 Classifications as a Policy Instrument of Inclusion and Exclusion

At the turn of the last century, classifications of feeble-minded people were assumed by different social reformers, worried by the rise of poverty, insalubrity and insecurity created by urbanization and massive immigration (Trent 1994). According to their discourses, feeble-minded people represented an important workload for the society and their increase required a political solution at a reasonable cost. Many of them shared the idea of creating colonies to group epileptics, morons, disabled and undisciplined people. Feeble-minded people had to be more productive. Policy-makers, concerned about mental deficiency, began to create specialized schools. With compulsory schooling, a new population was coming in schools and it challenged teachers who were claiming about indiscipline and delinquency. The management of deficient students became an increasing political concern whereas most experienced teachers were not capable of tackling with these “backward children” categorized as “silly, stupid, idiot, simple-minded, scatter brained, clogged, moron, duffer, dizzy, dull, peasant, uncultivated, airhead, squash, etc.”

From a sociological perspective, classifications are linked to key cognitive operations including and excluding human beings (Popkewitz 2013). As it has been exemplified by Emile Durkheim and Marcel Mauss (2009), “primitive” and “scientific” classifications share a common nature: they make relationships between human beings intelligible. The social function of classifications corresponds to a cognitive order for accessing to knowledge. Classifications between things shared among individuals and groups help to understand the logic of most decisive categories for the human mind: space, time, causality, etc. Following these works, Mary Douglas (1986) explains that the design of classifications is a specific exercise of polarization and exclusion: it implies tracing boundaries and creating equivalences between things which are a priori not comparable. Classifications institutionalize a hierarchy which is not only cognitive but social with important consequences on structuring relationships and power within the society. Indeed, education credentials, diplomas, certifications in modern times have been the mean to classify indi-

viduals according to their knowledge and skills and to position them in the social hierarchy serving the political objectives of sorting and selecting people.

In transferring classification into the issue of social classes, Pierre Bourdieu (1989), by criticizing the realist and Marxist conception of production relationships, has formulated a theory of social fields in which the fight for classifications, particularly through the school system, functions as a mechanism of social reproduction legitimizing certain ranks, titles and hierarchies. Some subtle distinctions are established according to the ownership of economic, social and cultural capitals which determine the ranks of individuals and groups within the social space. From them, incorporated dispositions (tastes, desires, affinities, etc.) are organized which correspond to practices and habitus of a social class with its relational properties. In conceptualizing a model of differentiation based on power relationships, Bourdieu demonstrates that systems of classification are the product of permanent struggles which redefine borders and modes of legitimacy structuring hierarchy and ranks and serving a politics of inequalities. In education, the metrics of inequalities has been extended to international surveys based on other modes of classification of equity and performance disconnected from social class theory.

16.2.2 Global Metrological Policy and Classifications of Inequalities in Education

Classifications have certain stability in time and space. Beyond the symbolic consecration and legitimation of differences, this representation of social order corresponds to a social and political investment. Statistical classifications play a dominant role in the legitimation of politics in education (Thévenot 2011). They guarantee three types of representation: a scientific and technical representation by which statistical tools allow to build and display a simplified reduction of society through charts and graphs. And it is also a political representation in the sense that social actors fight and negotiate for being represented and for representing their interests within the classification. Geoffrey C. Bowker et Susan L. Star (2000) show, from the international classification of diseases, that classification is the result of a compromise between several interests related to national and local systems of information in public health policies. Classification serves also as a cognitive representation and mental picture of social reality which allows to identify ourselves and those with whom we developed relationships.

From this perspective, it is easy to understand what is at stake in the definition of international nomenclatures in education as a global policy. Nomenclatures impose a universal system of classification even if their apparent homogeneity is questionable, according to historical, social and cultural differences between countries. However, The UNESCO's International Standard Classification of Education (ISCED) or the regular publication of the OECD's Education at Glance is rarely

challenged from a methodological and political stance. These classifications operate as “black boxes” for which data are legitimized by prestigious institutions and experts without questioning their degree of “harmonization” (Normand 2009). Beyond a realistic vision validating the measurement according to its biased or not biased dimensions, a sociological and constructivist perspective has shown that this policy of measurement itself depends on procedures and comparisons conceptualized according to certain rules of observation, recording and coding. It is the result of a complex networking and material assemblages between human beings and objects particularly when it requires the translation of a language in another or the conversion of “indigenous” cognitive categories into “universal” ones (Gorur 2011, 2014).

What is at stake in these modes of classification can be analysed through the debate on the PISA survey measuring equity and performance. Behind the “grey zone” of international surveys, difference agencies and transnational experts define categories of knowledge and thoughts which are transferred in time and space (Pettersson et al. 2017). This belief in numbers is underpinned by powerful technologies of calculation and entrepreneurial logics promoting “best practices” for education policy. At international and national levels, PISA acts as a “boundary object” opening the field to moral and political entrepreneurs who use the results and media to advocate their representations and interests in the public space challenging the current state of education systems (Normand 2014). Scientists and experts use these classifications to confront their arguments about the means of improving the effectiveness of education systems and developing accountability and New Public Management policies. From these data, journalists and some intellectuals seek to arbitrate their ideological quarrels between doxa and philosophy to influence policy-making. Policy-makers, converted to new forms of a pseudo-scientific experimentalism, find some ideas of rationalization and justification of their unpopular reforms. Some international agencies and experts produce tools, reports and recommendations for them, and they organize peer learning activities and exchange of best practices influencing national policies.

Among these classifications in education, the new born is benchmarking (Bruno 2017). At the beginning, it is a managerial technology implemented by the company Rank Xerox. It has quickly penetrated public policies to guarantee a process of objectivation of “best practices” in developing comparisons of performance and justifying decision-making. Assigning objectives to an indisputable realism, it promotes an art of government by probing data which subordinates public policy to a process of voluntary deliberation without hierarchy and rules. Agents of this policy are engaged in exchanges and debate on facts and numbers which require implicitly a convergence. That is why the European Commission has used benchmarking as an instrument of the Open Method of Coordination for education (Lange and Nafsika 2007). It is a “soft governance” not imposed to the States but leading them to consider their respective rankings for improving their equity and performance under fixed and precise targets. This technology of benchmarking is today used by influent

consultancy groups, like Mc Kinsey, to classify education systems considered as the most performative and equitable and to address recommendations to policy-makers (Gunter and Mills 2016).

16.3 Experiments as Policy Instruments

Experiments come from medicine. By breaking with the principle of dissection of corpses which allowed anatomy to become a science, Claude Bernard proposed a counterfactual experimental approach to highlight functions and symptoms of the human body. Moreover, the experimental method has strongly renewed clinical medicine in giving importance to the laboratory while it has also influenced health policies. Since, experiments on human beings were developed, first on convicts, then on prisoners and disabled people, before campaigns of vaccination on children (Lederer 1995). Then, psychology, inspired by medicine, sought to promote experiments and tests on human behaviour in the field of education politics, not without introducing some eugenicist ideas. The modern policy of experiments has inherited from these metrics to rehabilitate an experimentalism converted into Evidence-Based Education.

16.3.1 *From the Laboratory Study to Eugenics*

In emphasizing experiments and quantitative methods, such as medicine and physiology, psychology determined a division of labour between experimental subjects, the source of producing data, and experimenters who manipulated the conditions of the experiment (Danziger 1994). This method was in competition with another model; clinical psychology in which patients were assessed as “subjects” compared to the performance of other individuals regarded as “normal” or “abnormal”. Clinical psychology aimed at measuring the impact of a particular or abnormal characteristic of a subject, according to his personality, while experimental psychology claimed to set up a universalistic process related to all human minds.

These two methods were different from the one imagined by Francis Galton (Godin 2007). The British psychologist, founder of eugenics, had settled a laboratory in London for testing “mental faculties” among individuals. They were chosen among ordinary people. A map with the inventory of their mental capacities was delivered for 3 pence. The psychologist aimed to build a data bank on human capacities to provide recommendations in terms of social, rational and effective planning of the population. If clinical practice or experimental psychology was concerned by analysing individual processes, Galton and his eugenicist followers, wanted to include experimental data into statistical series to produce metrics of performance at large scale and to facilitate decision-making for social policies, including health and education (Bashford 2007; Lowe 1998).

To meet the needs of educational administrators and policy-makers, psychologists developed different methods: firstly, the experiment in laboratory but it was maladjusted for large-scale studies. Secondly, mental tests which allowed to compare individual differences through statistical series. They gave the opportunity to set up performance standards and categories from which individuals could be ranked according to criteria from the “general intelligence” of eugenics to the required qualities of a “good seller” (Kevles 1985). A third technology, the experiment in classrooms, offered new possibilities for psychologists. It facilitated the study of a group of students exposed to different methods of instruction while their performance was assessed before and after the experiment. Experiments in classrooms allowed to compare the efficiency of different techniques of learning and instruction whereas mental testing allowed to select individuals for adjusted social programs. They were constantly expanded during the 1920–1930s.

The abandon of the individualistic perspective in the collection of psychological data was linked to the building of a statistical rationale and more demanding modes of totalization for governing population (Ramsden 2003; Soloway 2014). The objective was to overcome traditional methods of comparing averages and ratios on population provided by statistical studies. While the dominant psychology relied on the model of experiment, statistical surveys on human conduct tried to study crime, suicide, poverty and health outside the laboratory serving the needs of the emergent and new Welfare State. Statistical societies compiled and analysed data to inform social reformers in guaranteeing them a better “scientific” approach. The study on the work conditions of children could be reinvested in the study of schooling whereas the production of school statistics increased (Travers 1983). The use of statistical charts allowed to reduce social problems to “objective facts”, to locate regularities behind variations of statistical numbers and to explain some mental behaviours. Human conduct was subjected to scientific and quantitative laws which would help psychological science to compile and combine more and more data.

Beyond developing tests to select talents, some ideas were shared in the USA among intellectuals that race and heredity play a fundamental role in the human development. Eugenicians were assuming and advocating restrictive immigration and segregation policies against those they judged unfit. Supporting selective reproduction programmes, they were influent on courts and local authorities while they were requiring sterilization and the dissemination of eugenicist ideas in textbooks and testing practices in schools (Selden 1999). In the UK, eugenicians studied the links between demography and degeneracy and they gave legitimacy to metrics on the quantity and quality of the population (Soloway 1990). It has important consequences in the areas of health and education policies while the Welfare State was elaborating its institutional and legislative frameworks. Issues of protection against diseases, replacement of the working generation, improvement in human capital, fight against waste were discussed whereas new ideas were emerging on economic efficiency and planning, redistribution and social justice in education as well as in other public areas. The *London Schools of Economics* was active in spreading these new conceptions

among scientists, intellectuals and policy-makers and in inventing instruments for a new political arithmetic of inequalities extending social and demographic surveys to educational issues (Normand 2011).

16.3.2 The Emergency Policy of “Controlled Experiments” and Evidence

If experts of efficiency expected psychological research developing metrics and comparing student achievement, they also required assessing the impact of different types of policy intervention (Danziger 1994). For the latter, psychologists had to compare groups of students exposed to different programmes (Sharp and Bray 1980). They were subjected to different experimental conditions with measurements taken before and after the intervention. The possibility of exploring the impacts of varied conditions of work within classrooms from outcomes measured by tests was a powerful motive for linking statistical data to experiments. Some studies on groups of treatment emerged in specialized journals, a lot of them dealing mainly with issues of fatigue and learning among students. The “controlled experiment” became a reference for comparing the efficiency of different administrative and political interventions.

Although the official story of “controlled randomized trials” begins with the Fisher’s experiment in agriculture, this technology was adopted before the 1930 in the field of psychology (Dehue 2001). Progressively, the experimental approach was extended to US educational research (Travers 1983). Psychologists, passing contracts with administrators and policy-makers concerned by efficiency, left their laboratories for experiments within local school systems. “Treatments” focused on teaching methods, discipline and punishments, and every teaching and learning behaviours in the classroom. If the *American Journal of Psychology* did not mention much controlled experiments during this period, 14% of the *Journal of Educational Psychology*’s articles were using the method (Danziger 1994).

One of these experiments was led by Thorndike and his student William Anderson McCall. The objective was to randomly assess impacts of fresh or regenerated air on student achievement measured by mental tests. McCall exposed the method in his textbook titled *How to Experiment in Education* (1923). He justified this type of experiments by the economy of dollars for school management. The book presents different methods of controlled experiments and randomization before the publication of Ronald A. Fisher which became a classic (Fisher 1925). But, despite its promising beginnings, the methodology of treatment group and controlled experiment took time to be developed. A certain pessimism surrounded this approach and it suffered from the loss of influence of the movement for efficiency during the 1930s.

Controlled experiments got a new legitimacy during the 1970s while social intervention programmes were discarded in the USA and issues of experiment were coming back. The intervention of the Federal State in social policies and compensa-

tory education programmes was criticized and the diminution of federal expenditures forced public authorities to adopt more short term and narrow interventions. Previous evaluations of these social programmes had been disputed as well as their methodologies (Cook 2000). It provides windows of opportunity for experts advocating new methodologies in metrics. It was the case of Donald T. Campbell who had published in 1969, an influent paper, which was a call for the USA and other nations to adopt “a new experimental approach of social reforms” based on specific treatments of social problems (Campbell 1969). The “true” experiments implied groups of individuals subjected to a treatment and compared to a control group. And if possible, the evaluation of the public policy, to be validated, had to overcome humanitarian and practical objections to expose randomly individuals to treatments during the time of the experiment. It was only for opposed moral reasons that other devices or statistical techniques had to be implemented. “Reforms as experiments” were not the first publication of Campbell, taking McCall as a reference, and he had earlier advocated the idea of extending the “logic of laboratory” to the society. With the statistician Julian C. Stanley, he had published a long chapter titled “experimental and quasi-experimental devices for research on teaching” (Campbell and Stanley 1963). In 1966, this chapter was republished in another book titled *Experimental and Quasi-Experimental Designs for Research* (Campbell and Stanley 1963). This last book was a best-seller promoting a new “standard” for research in social sciences and considering each researcher as the “methodological servant of the experimental society”. In the USA, afterwards, controlled randomized experiments became the “true experiment” and numerous public policies were implemented in education, health and social work following these principles and criteria.

Although the Federal State’s action diminished during the 1970–1980s, Campbell’s ideas were resumed in education by a Right-wing coalition for evidence-based policy who imposed these technologies in US education through an important lobbying beside the Congress (Normand 2016). Inspired by methodologies used in medicine (controlled randomized trials, meta-analyses, systematic reviews of research literature), experiments became a standard for the No Child Left Behind policies (2001) while its principles were resumed by International Organizations and exported in Europe. Evidence-Based Education have been since a reference for policy-making but also New Public Management (Wells 2007). The postulate of developing educational research and practices on “what works” entailed the creation of specialized agencies and international consortiums, such as the Campbell International Collaboration, to produce an influent expertise for policy-makers and putting pressure on researchers and practitioners (Lingard 2013; Trimmer 2016). Controlled randomized trials, largely advocated by economists in education, are today regarded as the “golden rule” for the evaluation of social policies, including education, and for the care of people qualified “at risk”. Controlled experiments and classifications in target-groups became the two pillars of the neo-liberal State’s new modes of social intervention which renounced progressively to universalistic mechanisms of allocation in making individuals accountable for their own behaviours through New Public Management techniques (Cribb and Gewirtz 2012).

16.4 The Politics of Standardization

Standardization allows to build uniformity in time and space in creating common standards and establishing political control on work and communities of practice at a distance (Brunsson et al. 2000). It helps the State and public authorities to compare individuals and groups and to adopt a common language shared by professionals, policy-makers and evaluators. Standards assume a mode of classification and measurement which defines limitations and exclusions in shaping a new policy. They lay on scientific and/or expert conventions and knowledge giving them legitimacy (Busch 2011). Their technicity prevents any reconsideration and controversy particularly when they result in a strong mobilization of expertise in time and space. Indeed, standardization is a policy instrument of power and coercion which effectively replace traditional rules of authority and hierarchy. That is why standards are often claimed on behalf of modernization and modernity which, in overcoming previous regulations, promotes a new Reason. For understanding the foundation and developments of standardization as politics in education, is it useful to consider the US history without forgetting that standards are today globalized through international surveys, the development of assurance-quality mechanisms in education and the promotion of “World Class” schools.

16.4.1 *Local Policies, Management of Efficiency and Standardization*

Above, we have already shown that in the USA, from 1880s to 1930s, new administrators and policy-makers shared a common expertise and belief in managerial effectiveness thinking that science, based on the systematic collection of data, would be able to create a new educative local and political order (Tyack 1974; Tyack and Hansot 1982). The time of the 3 R (Writing, Reading, Arithmetic) was achieved. All students, according to their innate talents, would be able to acquire standardized knowledge for their success in public education. Administrators wished to promote a new policy based on transparent standards, stratified and hierarchical school organizations, objective criteria to value individual skills. The politics of standardization on behalf of efficiency had to be underpinned by academic research and methods coming from the industrial world.

At that time, the US education debate was split between the Ancients and the Moderns (Cremin 1964). On one side, the generation of Horace Mann and the partisans of the Common School wanted education policy to consolidate the school system on a moral basis in emphasizing civic principles, communitarian consensus and local democracy. On the other side, professionals of management, qualified later as “progressive administrators” or “educative trust”, thought that education policy could be regulated by instruments of scientific progress and expertise guided by the production of standards. They expected “getting out politics from schools” in

subjecting school organization to new engineering (Tyack 1974). In adopting the model of the Taylorian company, school boards, including representatives from different communities, would be replaced by superintendents and managers concerned by effectiveness and the fight against waste in management.

In addition to their proximity with psychologists, progressive administrators were inspired by the scientific management implemented in big industrial companies (Callahan 1962). An effective manager had to collect the maximum of quantitative data to set up policy standards. The purpose was to better know the number of students in each school district, the number of school buildings, test scores, etc. Budgets had to be justified in terms of cost-effective methods. Progressively, these managers and experts imposed their political perspectives on effectiveness and standardization for school curricula. One of their eminent spokesmen, John Franklin Bobbitt, advocated a curriculum policy based on the measurement of efficiency and standards (Callahan 1962). He defined a scientific conception of the curriculum to improve school efficiency and to limit waste. The purpose was to decompose school subjects in precise objectives, then to split them in small units to improve the return of learning and teaching. This policy of standards was resumed by a lot of reformers who were also using psychological research on mental testing.

Considering the developments of the management of efficiency, it is easy to draw parallels with the current New Public Management. They share similarities in providing new opportunities for experts and policy-makers, changing the relationships with local and national authorities and converting professions to new ways of thinking and being accountable through standards (Gunter et al. 2016). They both use Taylorian mechanisms (ex. Quality assurance procedures) and incentives (ex. Performance related-pay) to put pressure and surveillance on educators (Ball 2003). The quest for limiting waste and adopting cost-effective measures is the same even if metrics have been modernized with the development of digital technologies. Rewards and sanctions, according to the meeting of objectives, are constantly a mean to achieve the 3Es: Economy, Efficiency, Effectiveness. What is probably new are the instruments of privatization (contracts, Public-Private Partnerships, outsourcing, etc.) which contribute to weaken and dismantle the Welfare State and the legacy of public authorities (Verger et al. 2016).

16.4.2 Towards an International Policy of Standards and Skills

Even if the movement for efficiency disappeared with the Second World War, the USA sought to maintain their quest for standards. Ralph Tyler, who was one of the psychologists converting IQ tests in knowledge and skills tests was at the root of an attempt of standardization and comparison of student knowledge during the 1960s (Finder 2004). The Kennedy-Johnson administration asked him to develop metrics on poverty in education. From 1964 to 1968, the ECAPE project (*Exploratory Committee on Assessing the Progress of Education*) gathered congress members, interest groups (notably Carnegie and Ford Foundations), representatives of US

States to design and develop the first federal assessment policy based on standards in school curricula (Lehmann 2004). Tests had to cover reading, English, mathematics and sciences, to diagnose strengths and weaknesses in the US education system. In fact, policy-makers were worried by the decline of standards in high schools after the launching of the Soviet Sputnik. It was urgent to train gifted scientists and engineers and to be more rigorous and demanding for curricula in sciences and mathematics. In 1968, the provisory committed became the NAEP (*National Assessment of Educational Progress*) and the first assessments of students was launched.

But there was pressure from the States to limit the extent of federal policy and the use of data was restricted as well as the follow-up of student progress. It is only after the publication of the report *A Nation at Risk* (1983) that the federal government paid attention again to the NAEP which was not producing any comparison between states. Its political and technical structure was completely revised and the US congress appointed a committee (the NAGB: *National Assessment Governing Board*) to develop standards on school achievement, to design tests, to publish scores and to ensure their dissemination at federal level. Since, NAEP assessments has become the benchmarking policy of US students' achievement particularly after the No Child Left Behind Act (2001) (Hursh 2007). The NAGB benefited from the expertise of the Education Testing Service, an agency specialized in the design of tests, created at its beginning by the US Navy to redefine the SAT (the test of entrance for prestigious US colleges) (Lehman 2001).

During the 1980s, while US political pressure on OECD was enhanced to develop and extend international surveys, the NAEP served as a reference for revising the first IEA surveys on mathematics. The *Assessment of Mathematics and Sciences* (IAEP) reused the NAEP items whereas the Education Testing Service imposed progressively its expertise for designing the PISA project. The survey was achieved during three cycles in 2000, 2003 and 2006 aiming to measure student skills at age 15 in reading, mathematics, and sciences. While PISA was resuming the methodological components of the NAEP, the IEA, and the ETS created a consortium (the IERI or *IEA-ETS Research Institute*) to develop research and analyses from international surveys, to train researchers and experts in these issues, and to disseminate standards worldwide. This policy of standards is today disaggregated at school level, with the survey PISA for Schools, and data serve to recommend best practices to potentially failing schools or to those who want to reach a global rank (Lewis 2017).

16.5 Conclusion

We have characterized three concomitating operations in metrics for education policy. Classification, by bringing things closer and ordering the world, make educative facts intelligible while it builds a truth of representation which shapes and guides

politics, particularly from knowledge produced by statistics and the collection of data. Experiments, in leaving laboratory and developing itself at large scale, allow to build statistical series used by experimental psychology and economics to qualify and classify populations according to different features and variables, and to prepare Post-Welfare State politics. While medicine serves as a reference, in education as well as in other areas of social policy, randomized controlled trials legitimize experiments as a cardinal principle positioning it above other methodologies used to produce knowledge. Metrics serve for building large bank of data on “what work” from which algorithmic treatments are considered as sufficient to establish evidence-based reformist proposals. Standardization is a policy by which, from metrics, the universe of practices is harmonized and subjected to standards or “best practices” denying cultural and contextual differences.

If metrics as politics is born with the development of administration of education and it concerns for efficiency, they are also technologies for governing school populations at large scale. In a time of globalization, New Public Management has adjusted Taylorian tools in modernizing them, experimental economics and cognitive sciences have discarded eugenicist assumptions from bio-economics and mental testing psychology, but a same rationalist and scientist temptation remains as it is shown by the success met by evidence-based policies. The process of instrumental Reason, as proved by this short history of the present, is a permanent quest for objectivity and truth through political claims which are constantly disclaimed by the irreducibility of human nature numbers or data (Biesta 2007). In this doomed attempt of reducing contingency and uncertainty to metrics, education politics paradoxically underpins sciences of government which narrows the range of possibilities for action and the plurality of individual and collective choices (Thévenot 2007). Subjected to control and reinforced surveillance by the sophistication of assumed perfectible tools, human beings must confront their potentialities and capacities to what is measurable at the expense of sacrificing their autonomy and self-fulfillment. This subjection of education to the government by numbers limits also the possibility to consider other forms of moral agency beyond figures of individual responsibility and expressions of competitive choice. At the end, it seems these John Dewey’s ideas have been completely forgotten by the apologists of metrics:

(...) moral equality cannot be conceived on the basis of legal, political and economic arrangements. For all of these are bound to be classificatory; to be concerned with uniformities and statistical averages. Moral equality means incommensurability, the inapplicability of common and quantitative standards. It means intrinsic qualities which require unique opportunities and differential manifestation; superiority in finding a specific work to do, not in power for attaining ends common to a class of competitors, which is bound to result in putting a premium on mastery of others. Our best, almost our only, models of this kind of activity are found in art and science. There are indeed minor poets and painters and musicians. But the real standard of art is not comparative, but qualitative. Art is not greater and less, it is good or bad, sincere or spurious. Not many intellectual workers are called to be Aristotles or Newtons or Pasteurs or Einsteins. But every honest piece of inquiry is distinctive, individualized; it has its own incommensurable quality and performs its own unique service. (Dewey 1922)

References

- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228.
- Bashford, A. (2007). Nation, empire, globe: The spaces of population debate in the interwar years. *Comparative Studies in Society and History*, 49, 170–201.
- Berliner, D. C., & Biddle, B. J. (1995). *The manufactured crisis: Myth, fraud, and the attack on America's public schools*. New York: Longman Publishers.
- Biesta, G. (2007). Why “what works” won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57(1), 1–22.
- Bourdieu, P. (1989). Social space and symbolic power. *Sociological Theory*, 7(1), 14–25.
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge: MIT Press.
- Bruno, I. (2017). “Silencing the disbelievers”. Games of truth and power struggles around fact-based management. In R. Normand & J.-L. Derouet (Eds.), *A European Politics of Education. Perspectives from sociology, policy studies and politics* (pp. 140–154). London: Routledge.
- Brunsson, N., Jacobsson, B., et al. (2000). The contemporary expansion of standardization. In N. Brunsson, B. Jacobsson, et al. (Eds.), *A world of standards* (pp. 127–137). Oxford: OUP.
- Busch, L. (2011). *Standards: Recipes for reality*. Cambridge: MIT Press.
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, 24, 409–429.
- Campbell, D. T., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Callahan R. (1962). *Education and the Cult of Efficiency*. Chicago: University of Chicago Press.
- Cook, T. D. (2000). The false choice between theory-based evaluation and experimentalism. In P. J. Rogers, A. Petrosino, T. A. Heubner, & T. A. Hasci (Eds.), *Program theory evaluation: Practice, promise, and problems, new directions in evaluation* (Vol. 87, pp. 27–34).
- Cremin, L. A. (1964). *The transformation of the school: Progressivism in American education 1876-1957*. New York: Vintage Books.
- Cribb, A., & Gewirtz, S. (2012). New welfare ethics and the remaking of moral identities in an era of user involvement. *Globalisation, Societies and Education*, 10(4), 507–517.
- Danziger, K. (1994). *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press.
- Dehue, T. (2001). Establishing the experimenting society: The historical origin of social experimentation according to the randomized controlled design. *The American Journal of Psychology*, 114(2), 283.
- Dewey, J. (1922). Individuality, equality, and superiority. *New Republic*, 33, 61–62.
- Douglas, M. (1986). *How institutions think*. Syracuse: Syracuse University Press.
- Durkheim, E., & Mauss, M. (2009). *Primitive classification (Routledge revivals)*. London: Routledge.
- Finder, M. (2004). *Educating America: How Ralph W. Tyler taught America to teach*. Westport: Praeger.
- Fisher, R. A. (1925). *Statistical methods for research workers*. London: Oliver & Boyd.
- Foucault, M. (1977). *Discipline and punish*. Harmondsworth: Penguin.
- Foucault, M. (2002). *The order of things: An archaeology of the human sciences*. New York: Psychology Press.
- Foucault, M. (2012). *The archaeology of knowledge*. New York: Vintage.
- Godin, B. (2007). From eugenics to scientometrics: Galton, Cattell, and men of science. *Social Studies of Science*, 37(5), 691–728.
- Gorur, R. (2011). Policy as assemblage. *European Educational Research Journal*, 10(4), 611–622.
- Gorur, R. (2014). Towards a sociology of measurement in education policy. *European Educational Research Journal*, 13(1), 58–72.
- Grek, S. (2009). Governing by numbers: The PISA ‘effect’ in Europe. *Journal of Education Policy*, 24(1), 23–37.

- Grek, S. (2013). Expert moves: International comparative testing and the rise of expertocracy. *Journal of Education Policy*, 28(5), 695–709.
- Gunter, H., & Mills, C. (2016). *Consultants and consultancy, the case for education*. Dordrecht: Springer.
- Gunter, H. M., Grimaldi, E., Hall, D., & Serpieri, R. (Eds.). (2016). *New public management and the reform of education: European lessons for policy and practice*. London: Routledge.
- Hood, C. (1986). *The tools of government*. Chatham: Chatham House.
- Hursh, D. (2007). Assessing no child left behind and the rise of neoliberal education policies. *American Educational Research Journal*, 44(3), 493–518.
- Keyes, D. J. (1985). *In the name of eugenics: Genetics and the uses of human heredity*. No. 95. Cambridge: Harvard University Press.
- Lange, B., & Nafsika, A. (2007). New forms of European Union governance in the education sector? A preliminary analysis of the open method of coordination. *European Educational Research Journal*, 6(4), 321–335.
- Lascoumes, P., & Le Galès, P. (2007). From the nature of instruments to the sociology of public policy instrumentation. *Governance Understanding Public Policy Through Its Instruments*, 20(1), 1–21.
- Lawn, M. (Ed.). (2013a). *The rise of data in education systems: Collection, visualization and use*. Oxford: Symposium Books Ltd.
- Lawn, M. (2013b). Voyages of measurement in education in the Twentieth Century: Experts, tools and centres. *European Educational Research Journal*, 12(1), 108–119.
- Lawn, M., & Normand, R. (Eds.). (2014). *Shaping of European education: Interdisciplinary approaches*. London: Routledge.
- Lederer, S. E. (1995). *Subjected to science: Human experimentation in America before the second world war*. Baltimore: Johns Hopkins University Press.
- Lehman, N. (2001). *The big test. The secret history of the American meritocracy*. New York: Farrar, Straus & Giroux.
- Lehmann, I. J. (2004). The genesis of NAEP. In L. V. Jones & O. Olkin (Eds.), *The nation's report card. Evolution and perspectives* (pp. 25–92). Bloomington: Phi Delta Kappa Educational Foundation.
- Lewis, S. (2017). Governing schooling through ‘what works’: The OECD’s PISA for Schools. *Journal of Education Policy*, 32(3), 281–302.
- Lingard, B. (2013). The impact of research on education policy in an era of evidence-based policy. *Critical Studies in Education*, 54(2), 113–131.
- Lowe, R. (1998). The educational impact of the eugenics movement. *International Journal of Educational Research*, 27(8), 647–660.
- McCall, A. W. (1923). *How to experiment in education*. New York: MacMillan.
- Miller, P., & Rose, N. (2008). *Governing the present*. Cambridge: Polity Press.
- Normand, R. (2009). Expert measurement in the government of lifelong learning. In E. Mangenot & J. Rowell (Coord.), *What Europe constructs: New sociological perspectives in European studies* (pp. 225–242). Manchester: Manchester University Press.
- Normand, R. (2011). *Gouverner la réussite scolaire. Une arithmétique politique des inégalités*. Berne: Peter Lang, Presses de l’École Normale Supérieure.
- Normand, R. (2014). The French pinnacle of PISA. In M. Lawn & R. Normand (Eds.), *Shaping of European education. Interdisciplinary approaches* (pp. 32–49). London.
- Normand, R. (2016). ‘What works?’ The shaping of the European politics of evidence. In *Towards a New Europeanus Homo Academicus? The changing epistemic governance of European education* (pp. 95–125). Springer International Publishing.
- Pettersson, D., Popkewitz, T. S., & Lindblad, S. (2017). In the grey zone: Large-scale assessment-based activities betwixt and between policy, research and practice. *Nordic Journal of Studies in Educational Policy*, 3(1), 29–41.
- Popkewitz, T. S. (1997). A changing terrain of knowledge and power: A social epistemology of educational research. *Educational Researcher*, 26(9), 18–29.

- Popkewitz, T. S. (2012). *Cosmopolitanism and the age of school reform: Science, education, and making society by making the child*. New York: Routledge.
- Popkewitz, T. (2013). The sociology of education as the history of the present: Fabrication, difference and abjection. *Discourse: Studies in the Cultural Politics of Education*, 34(3), 439–456.
- Porter, T. M. (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
- Ramsden, E. (2003). Social demography and eugenics in the interwar United States. *Population and Development Review*, 29(4), 547–593.
- Selden, S. (1999). *Inheriting shame: The story of eugenics and racism in America*. New York: Teachers College Press.
- Sharp, S. A., & Bray, A. P. (1980). WH Winch: A founder of the experimental approach in education. *British Journal of Educational Studies*, 28(1), 34–45.
- Soloway, R. A. (1990). *Demography and degeneration: Eugenics and the declining birthrate in twentieth-century Britain*. Chapel Hill: University of North Carolina Press.
- Soloway, R. A. (2014). *Demography and degeneration: Eugenics and the declining birthrate in twentieth-century Britain*. Chapel Hill: UNC Press Books.
- Thévenot, L. (2007). The plurality of cognitive formats and engagements: Moving between the familiar and the public. *European Journal of Social Theory*, 10(3), 409–423.
- Thévenot, L. (2011). Conventions for measuring and questioning policies. The case of 50 years of policy evaluations through a statistical survey. *Historical Social Research/Historische Sozialforschung*, 36(4), 192–217.
- Travers, R. M. W. (1983). *How research has changed American schools: A history from 1840 to the Present*. Kalamazoo: Mythos Press.
- Trent, J. W. (1994). *Inventing the feeble mind: A history of mental retardation in the United States*. Berkeley: University of California Press.
- Trimmer, K. (Ed.). (2016). *Political pressures on educational and social research: International perspectives*. London: Routledge.
- Tyack, D. (1974). *The one best system: A history of American urban education*. Cambridge: Harvard University Press.
- Tyack, D., & Hansot, H. (1982). *Managers of virtue: Public school leadership in America, 1820–1980*. New York: Basic Books.
- Verger, A., Fontdevila, C., & Zancajo, A. (2016). *The privatization of education: A political economy of global education reform*. New York: Teachers College Press.
- Wells, P. (2007). New labour and evidence based policy making, 1997–2007. *People, Place & Policy*, 1(1), 22–29.

Romuald Normand is professor of Sociology at the University of Strasbourg, Faculty of Social Sciences, France. (Research Unit CNRS SAGE: Societies, Actors and Government of Europe) and Associate Visiting Professor and Head of the Chinese-French Centre for Innovation in Education (CFCIE), Beijing Normal University. Professor Normand works on comparative education policies and politics, Europeanization and Lifelong Learning, Higher Education and Research, New Public Management in education. Romuald Normand is convenor of the network 28 “Sociologies of European Education” at the European Educational Research Association. He is a member of the editorial board of the British Journal of Sociology of Education and co-editor of the Routledge series “Studies in European Education” with Martin Lawn.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.