



# ECNU Review of Education

ISSN: 2096-5311

CN: 31-2150/G4

<http://www.roe.ecnu.edu.cn>

---

## **Calculating the Future: The Historical Assemblage of Empirical Evidence, Benchmarks & PISA**

DOI 10.30926/ecnuroe2018010106

Thomas S. Popkewitz, Jingying Feng, and Lei Zheng  
University of Wisconsin-Madison



## Calculating the Future: The Historical Assemblage of Empirical Evidence, Benchmarks & PISA<sup>1,2</sup>

Thomas S. Popkewitz, Jingying Feng, and Lei Zheng  
University of Wisconsin-Madison

© East China Normal University & East China Normal University Press, Shanghai, China

### Abstract

**Purpose**—Prominent at the intersections of national educational agencies, higher education, and international educational performance assessments are two reform standards: “benchmarks” determining optimal student performance, and “empirical evidence” for determining the quality of reform practices. These two notions are often taken as connecting policy and research to effective changes in many countries. The article examines the historical and cultural principles about educational change and its sciences embedded in these standards through examining OECD’s PISA and the McKinsey & Company reports that draw on PISA’s data.

**Findings/Originality/Value**—First, the reports express salvation themes associated with modernity; that is, the promise of a better future through governing the present. The promise is to provide nations with data and models to achieve social equality, economic prosperity, and a participatory democracy. Second, the promise of the future is not descriptive of some present reality but to fabricate the universal characteristics about society and individuals. The numbers embody social and psychological categories about a desired unity of all students. Third, the “empirical evidence” of the international assessment entails a particular notion of science and “evidence”; one that paradoxically uses the universals in comparing and creating divisions.

### Keywords

Empirical evidence; benchmarks; international assessment; system of reason; historicization

There is almost a truth to contemporary school reform. The truth is that to guarantee the quality of reform is to identify *benchmarks* as standards to achieve and to judge students and teacher success through research that provides “*empirical*” or “*scientific evidence*”. Benchmarks and empirical evidence are what makes possible the connection of policy and research to effective change. This notion of change is prominent at the intersections of national educational agencies, higher education, and comparative metrics of educational performance assessments in many countries, such as China and Sweden. The use of benchmarks and empirical evidence are assumed as foundational for models of change found in social services agencies and national educational systems.

The notions of benchmarks and having “empirical evidence” to validate reforms, are not merely about science. When thought historically and culturally, they embody

salvation themes that promise a better future through governing the present. The promise is to provide nations with data and models to achieve social equality, economic prosperity, and a participatory democracy. This is expressed, for example, in international assessments such as OECD's PISA. This future is sometimes given names, such as "the Knowledge Society" and "the Innovative Society" in many international documents. The international ranking lists of universities and school systems, for example, are coupled with models of change that speak about the assessment reports as providing the data that can enable nations to have the world's best-performing school systems (Mourshed, Chijioke, & Barber, 2010). Benchmarks are the technologies to optimize the qualities and characteristics for the nation to function efficiently and achieve prosperity. The numbers in the assessments serve as guarantees that the true, efficient, and effective empirical evidence to chart national change is represented.

The salvation themes of the future are a particular kind of utopic thought that emerges in the long 19th century about science and social changes (Popkewitz, 2008; Wang, 2008). Utopic in the international assessments are not merely descriptive of school systems but are embodied models of change directed to the future. These models of change, as we argue, are about research as actualizing what people and society should be. This utopic quality is obscured through the benchmarks and empirical evidence that seem as universal qualities and characteristics associated only with school success. But the categories, algorithms and statistical computations express particular cultural principles or "theses" about the perfection of the child for the future through educational reform.

The salvation themes for thinking about and organizing national reforms are explored in two prominent policy oriented efforts to assess and organize changes in educational systems: PISA, an international survey which assesses worldwide student skills and knowledge in science, mathematics, and literacy and the McKinsey & Company educational reports, which draw on PISA results to "help educational systems and providers to improve outcomes for millions of students globally".

The ideas of benchmarks and "empirical evidence" in the assessments are particular historical notions rather than universal ideas of science. They embody a particular historical vision that is not merely descriptive but principles to be actualized through *making kinds of people*; that is, as determinate categories about the qualities and characteristics of populations (see, e.g., Popkewitz, 2008). Benchmarks and "empirical evidence" do not "merely" operate to describe the world for people to act on. They are actors in social affairs. Benchmarks and "empirical evidence" are assembled and connected in a particular historical mode of visualizing problems, its notions of methods, and what counts as solutions to social issues. To speak of this a little differently, benchmarks and empirical evidence are like a cake. They are given intelligibility with a set of ingredients that when brought together creates the objects of seeing and acting on as important for change.

This way of seeing and acting in research is explored as a system of reason; that is, an assemblage of different historical events that come together to order and

classify reflection and action. The notion of system of reason is a theoretical way of thinking historically about the principles that order and classify what is seen, acted on, and thought about in educational reforms. This notion of system is different from what we later consider as “systems theory”. That later is a particular historical set of rules and standards that underlies what is taken as the objects and relations about schools and achievement studied in the international assessments. If we can play with words, benchmarks and “empirical evidence” are explored as words performed within a particular set of rules that makes it possible to take the international assessment as reasonable and plausible to think about what occurs in schools. Methodologically, the approach is a history of the present. Our use of “system” in thinking about reason is a conceptual way to think about the historical principles that organize reforms, research and assessment as saying “see this”, notice this, by providing distinctions and classification to judge what is important to schools and assess change. This notion of system of reason is different from “system” theory that orders PISA’s measurements, the later is discussed as a particular way of classifying, organizing, and managing social affairs and populations.

### **How the Recipe of Benchmarks and “Empirical Evidence” Becomes Possible: The “Reason” of Systems**

We would like to discuss two historical dynamics in the making of the benchmarks and the ideas of “empirical evidence” before moving to the international assessments. One relates to the formation of social science in the long 19th century; that is, overlapping historical trajectories that come together and are institutionalized as the social and psychological sciences by the turn of the 20th century. The second are changes that occur in the social sciences after World War II through the joining of systems theory with cybernetics. This is not meant as an evolutionary history but a history of the continual assembly and connections that entail continuities as well as discontinuities.

#### ***Forming the Social Sciences, Making Kinds of People and Differences***

Finding the commonsense of benchmarks and what counts as “empirical evidence” historically is in the emergence of what was called initially “moral sciences”. This may sound odd as benchmarks and “empirical evidence” are thought of as neutral practices, descriptive practices outside of ideologies and social and moral value; they are thought of as only a descriptive knowledge about what works. Yet these phrases are not outside of human history but part of it. If we look to the beginning of the 1800s, the sciences about human conditions and people were called “moral sciences”. The concerns were with issues of deviancy and how to correct moral disorder by making kinds of people. This making of people embodied double gestures

of the Enlightenment. The gesture of hope was that through the applications of reason and rationality, pathways to progress would be found to bring liberty, prosperity, and happiness to humanity. But moving with the gestures of hope were fears of the dangers and the dangerous populations. The populations embodied threats to the desired futures; talked about in the 19th century as barbarians, savages, backward and today spoken about with other notions to differentiate and distinguish cultural and moral differences, such as immigrants, ethnic groups, the “at-risk” child, and “fragile” families as different from some unspoken normalcy and the “at-risk” child and “fragile” families.

Let us provide two examples of science and the making of kinds of people. One is the turn of the 20th century psychologies of pedagogy. One central figure was the American John Dewey (Popkewitz, 2005, 2010). Dewey argued that the democratic school and pragmatism was a moral philosophy for interpreting how Christian ethics and the moral good could be brought into social and educational affairs (Dewey, 1967–1990, originally published in 1892). Dewey brought this way of thinking about salvation theme into the relation of science, moral order, and pedagogical practices.

Science was not about what scientists did, but a social project to bring into being a mode of thinking and ordering change and uncertainty through “problem-solving”. The hope of democracy, as expressed by Dewey, was an anthropological psychology producing the future cosmopolitan child through method of observation, description, and induction. In the gesture of hope of cosmopolitanism were fears about the dangers and dangerous populations. At the turn of the century in America, these fears were embodied in the qualities of life associated with the poor, immigrants and racial groups of the new industrial cities. The dangers were also in the unbridled capitalism of this period where there were no restraints on industrialization and what people called “the Robber Barons”.

The name of moral sciences disappeared by the end of the century and was replaced with the new “disinterested” languages of science applied to the human conditions. The new language subsumed moral questions about the present and the future, however, into a seemingly more neutral and objective discourse that gave focus to the primacy of methods. According to Wang (2006), the discourse of a unified system of science that turned science into a foundation of ethics, aesthetics, and politics at the level of social practices also emerged in China around the 1920s. But to think about how science as making kinds of people is (re)visioned, reassembled and given the language that we now speak of as benchmarks and “scientific evidence”, the post-WWII years need to be brought into focus. This becomes the second part of the ingredients of the recipe that is assembled in the making of people.

A second example in the making of kinds of people is systems theory and cybernetics. Benchmarks and “scientific evidence” are given expressions through cybernetics to think about human affairs as an analogy of mind and machine. Initially tied to war efforts, cybernetics circulates as ways of thinking about cognitive psychology, “bounded rationalities”, political systems, sociological phenomena, and

anthropology. The systems theory and cybernetics created new set of axiomatic rules and standards to rationalize and administer social life in how it defined human nature and the mind in relation to the logic of machines. Brought into social analysis, cybernetics was created as a way to think about mind in relation to the machine—the machine as the computer and its analogy to the mind as artificial intelligence.

There is a paradox to the systems theory that is applied. The paradox occurs in concert with change related to the system's growth and development that, in actuality, take current frameworks about people and society as the starting point for understanding and potential reforms. The object of reforms and educational interventions is conceptualized in a manner that builds on what already exists in educational practices. This does not enunciate a theory of change. Change is nothing more than motion and activity. This conservatism in assumptions about the objects of change is hidden through the focus on methods about the processes and networks of communication as bringing about change. If recent scholarship in the history of science is examined, cybernetics provides concepts for mapping the processes and flows of information as stable objects for administration, the mode of reasoning whose principles give form to the current thinking of benchmarks and scientific evidence.

Cybernetics theories connect to systems thought. System is an abstraction to actualize future society and people; the abstraction embodies principles that are not empirically deduced but are *a priori* and *self-referential and self-authorizing*; that is, its mode of ordering and classifying inscribes internal boundaries in defining problems, contexts and the possibilities of change. If we think of cybernetics as particular epistemological rules and standards for constructing social phenomena, what we called a system of reason, this quality of being self-referential and self-authorizing is not unique. What is given focus here, however, are the principles of systems thought as a strategy of change.

The idea of a system as an organism replaced earlier mechanical notions with more dynamic models of change. But the idea of a machine did not disappear. With language borrowed from biology, social institutions are conceptualized as a *social organism* having stages of growth and processes of development that change over time.

There is a paradox in this reasoning about social relations and change. The paradox reoccurs as the idea of the school as a system is an abstraction about ontological objects that have machine-like qualities of growth, development and change which can be calculated and standardized to allow for what is not a machine. Cybernetics and systems thought move from the goal to obtain ideal types to thinking about standards concerned with optimizing utility of the system without striving for perfection. One of the debates in computer science during this time was whether the purpose of research was to create programs that eliminated all errors, thus producing the modern Philosopher's Stone. The other position was to try to produce programs that would eliminate errors as best as possible, knowing that the perfect system was not possible. This latter approach won!

The outcome in computer science was to reduce error as much as possible and thus bring to optimize the machine's efficiency as much as possible. This notion of optimizing performance is carried in the reform-oriented sciences and notions of benchmarks. To bring to the present, the international ranking systems of PISA and other social and economic indicators are optimizing the functions of school "system". The rankings draw on cybernetic modes of thinking to compare, order, and plan for efficiency to optimize and identify the "best" processes and communications patterns. That "best" practice, however, stabilizes "the system" and conserves its assumptions.

Another element in this new rationality was what constituted the rules and standards of "empirical evidence". Historically, the idea of "scientific and empirical evidence" means simply systematically observing what happens in everyday life. A newspaper, a play, a sport game, as well as introspection in early psychology were ways of ordering and classifying "empirical evidence". In post-war years, social science was concerned with the administration of change incorporating the idea of algorithms to think through mathematics about "empirical evidence". Algorithms, it needs to be noted, entails a particular kind of mathematical thinking about social life as having rigid rules that provide optimal solutions to given problems, or delineate the most efficient means toward certain given goals. The models of change offered by the OECD report on the Swedish school system (Pont, Donaldson, Elmore, & Kools, 2014), discussed later, inscribe the operation of algorithms as underlying principles for forming the model of change that is to lift Sweden from average to above average.

When cybernetics, systems theories, and "empirical evidence" are ordered as algorithmic rules, the numbers and benchmarks of international ranking become particular cultural practices about the making of society and people.

### ***Making Society/Making People: The Cultural Practice of Numbers***

By now, it should be clear that the benchmarks of international assessments of schools and international ranking of universities are not merely descriptions born of empirical data drawn from the present but historically embodied in trajectories of the social sciences that are about people to actualize a desired future. The OECD's PISA and the McKinsey reports on education are ordered through cybernetics and systems analysis as a theory ordering assessments by focusing on processes and communication patterns of social life that, while, at the same time, it is about ordering the possibilities of change that anticipate what is the desired future of an imagined society and people. The school is studied as a system that has qualities of a biological organism, a metaphor to think about "*the educational needs*" in which social growth and development can be measured.

Numbers serve as the reference within the systems analysis, and benchmarks as the empirical evidence. Numbers connect as a further ingredient of this recipe knowledge of assessment and change. The magnitudes of differences in the statistical

correlations are placed into models of intervention that are to bring into existence kinds of people that can actualize the effectiveness of school viewed through an abstraction of systems to think about and administer social relations.

If we move to the present and again being synoptic, international assessments of the OECD are not “merely” descriptive of some reality but “act” in making or fabricating what matters; what “acts” as a given solution to social problems and the strategies of change are to enact that “nature”. The statistics and numbers generated in the international assessments are taken as stable scientific facts for planning and interventions. Measures provide a comparative algorithm that “tells” of a continuum of values about people and the future that enables successful school systems. The continuum in the international assessments appear as stages or the scaling of educational systems and their “improvement” according to the OECD and McKinsey models (see, e.g., Mourshed et al., 2010).

The measures are to lead to a common world accessible as highways to rectify the dangers that are disruptive of the equilibrium of the system that is invented. That is what the models of change in the OECD Education Policy Review report of assessment and change are to produce. The models of change are not merely about systems. In the Swedish report, the universal characteristics and qualities of kinds of people are those that are actualized nationally, as the vision and rationality for thinking and acting as teachers, but also as the social and psychological qualities of “well-being” of the abstractions that unite students, parents and communities! (See, e.g., Pont, et al., 2014; OECD, 2017). Similarly, the McKinsey report *Drivers of student performance: Asia insights* takes China and Japan cases as a way addressing the qualities of performance. But performances are not merely descriptive of what students achieve in current school systems. Student performances are closely linked with their socio-economy, personality dispositions (called “mindsets”), and attitudinal attributes (Chen, Dorn, Krawitz, Lim, & Mourshed, 2017).

### ***Benchmarks & Variations: Desired People to be Actualized***

The counting and numbers when comparing nations and educational systems perform as expectations about universal characteristics of society and people whose composition forms a common and harmonious world. The numbers embody an anticipatory reasoning about the future society and populations. McKinsey's *How the world's most improved school systems keep getting better* argues, for example, that benchmarks are an “universal scale of calibration” to create equivalences from, for example, “several different international assessment scales of student outcomes discussed in education literature” (Mourshed et al., 2010, p. 7). Benchmarks are standards placed in scales that order elements on a continuum from “poor/fair to good”, “good to great” and from “great to excellent”. In a different report on how school systems are improving, the scale is given as a clear and linear progression that is internal to each category and then correlated across categories, such as:



*Fair to good:* consolidating system foundations, high quality performance data, teacher and school accountability, appropriate financing, organization structure, pedagogical models;

*Good to great:* teaching and school leadership as a full-fledged profession, necessary practice and career paths as in medicine and law; and

*Great to excellent:* more locus of improvement from center to school, peer-based learning, support of system-sponsored innovation and experimentation. (Mourshed et al. 2010, p. 20)

The strategy is to address deviations from the norms in the development of country case studies. Variations from the standardized norms are used to define differences and spaces of actions.

The benchmarks appear as referencing the nation in a continuum of development. But the statistical categories and their relations to describing development are about the qualities and characteristics of people. Chinese education inspection is an example. The Basic Education Quality Inspection Center of China started the project National Compulsory Education Quality Inspection (*Guojia yiwu jiaoyu zhiliang jiance*) in 2007. Its proposal was officially published in 2015. On its news release conference, Director of Center He Xiuchao saw the quality inspection as helping accelerate students' comprehensive development, promoting the quality of compulsory education, and leading the government and the general society to build up correct concepts of education quality (as cited in Ministry of Education, 2015).

The universal measures of achievement are correlated to who the teacher is, psychologies of the child, school organization, and norms about modes of living called "parent participation"; for example, "peer-led creativity and innovation" and "building technical skills of teachers and principals". Measurement categories that focus on "creativity", "innovation", and "participation skills" embody principles about desired kinds of people and the kind of society. These desires are organized as epistemological principles that compare students, families, and achievement levels to differentiate successful and less successful school systems on a continuum of value. The qualities and characteristics of the comparison embody unspoken values about differences. These differences are expressed as the psychological and social characteristics of students, such as about the "enjoyment of life", happiness, belonging, and self-realization that seem to have no historical/cultural location.

The logic of "well-being", as earlier, is not merely descriptive about children's lives. The distinctions and categories about children's differences produce a continuum of value. The differences are standardized, codified and ordered into universal hierarchies of values for comparing. The hierarchy of values is created to differentiate nations and populations. The statistical analyses used to talk about school systems are said to "examine why and what they have done have succeeded where so many others failed" (see, e.g., Mourshed et al., 2010).

The standardizing and codifying to find equivalences, ironically, erase difference

by establishing difference. The reduction of complexities to those of rational management “systems” makes it seem that “all” national systems can anticipate equality through the application of categories to recognize differences that inscribes difference. Differences entail comparisons through creating sets of equivalences among disparate databases. The paradox of the international comparisons is its inscription of difference that “makes” differences so that some can never be at the “top”.

### ***Double Gestures: The Hope and Fears of Kinds of People***

Benchmarks and their “empirical evidence” embody universals that paradoxically compare and create divides. Lists and rankings in the international assessments compare secondary statistical measures that create “*a universal calibration*” in which a spectrum of norms defines equivalencies among subsets of data (Mourshed, Farrell, & Barton, 2013). The comparison eliminates differences to produce distinctions that divide. If the OECD and McKinsey reports are examined, effective education travels as the gesture of hope that forecasts the salvation themes of a good society, full employment, well-being, and the progress of the nation. The classifications and numbers connect to psychological categories of children’s social and communicative patterns, such as family influence on children’s achievement and the relation of education to employment.

The social and psychological distinctions are about the hopes of future kinds of people. The hopes, however, simultaneously express the gesture of fear of the dangers and dangerous populations to that future. The fears are about the kind of child or parent who does not “fit”, is abjected and excluded for the characteristics and qualities related to school success. The measures, when placed in the history of psychology, are about moral order/disorder as when incorporated into models of change that tells of the abject child as “lacking” motivation, well-being, and the proper modes of living (Danziger, 1990; Cohen-Cole, 2014). The delineating of stages of development are not only organizational factors but they also align with psychological qualities of youth that normalize what is functional and dysfunctional for employability, such as being disengaged, disheartened, well-positioned or too poor to study (Mourshed et al., 2013, pp. 32–33).

The gestures of hope and fear are double gestures. The statistical calibrations are about who people are and should be, as well as about who do not “fit” as part of the universal. The characteristics of people who succeed and don’t succeed form a continuum of value. That continuum simultaneously embodies the hope to actualize a desired future with fears of populations inscribed as dangerous to the system’s harmony and consensus. Codifying and standardizing are not merely about achievement. The ranking and classification engender differences in those “civilized” and those different in degree from that advanced stage of civilization—the school systems and nations at the top!

### ***“Follow Me!” Knowing the Future as Taming Uncertainty***

The future is thought of as certain and the problem of measurement is to put nations and people on the highways to actualize the abstraction of the school system. McKinsey uses the highway metaphor, for example, to think about highways as not merely paths to the future. They embody the qualities and characteristics of the kinds of people who will inhabit that future. Not far away from the highways and pathways that are to “deliver better outcomes” for future harmony and consensus are fears. To follow the models of change in reducing unemployment among ethnic, racial and poor populations is as *“to get rid of potholes, make educators and employers part of the solution by providing ‘signs’ and “concentrate on the patch of pavement ahead”* (Mourshed et al., 2013, p. 54).

Benchmarks and “empirical evidence” are inscription devices that portray that the knowledge of the future is at hand for all nations to reach the top. The pathways posit social life as a machine whose proper alignment allows for it to administer system goals. The problem becomes simply how to tailor the highways individually so all can find the destination.

### **Some Concluding Thoughts**

Creating benchmarks and “scientific evidence” provides the contemporary temptations for finding the Medieval Philosopher’s Stone, that is, having the ultimate answer to issues of development and progress. The beckonings of today express salvation themes that are given specific enactments in the international assessments of student performance examined.

The discussion explored the particular limits in thinking about change in these contemporary international practices to assess and order educational systems. It was argued that change is directed by particular universal qualities and characteristics of people and society to be actualized through the research models. The anticipatory future embedded in the assessments and their models of change is a calculated rationality that appears as ahistorical and without any particular social and cultural values; yet the statistical categories and distinctions are produced in particular historical configurations of the social, psychological, and educational sciences related to student assessments. These historical inscriptions about people enabled us to talk about numbers as not merely descriptive but also cultural practices. Further, the anticipatory quality of the international assessments “acts” in the same manner as Google, Amazon, Baidu, Taobao, or Youku search algorithms that anticipate your wants and desires. The difference with the international assessments from the web searches is that our preferences have not been registered prior to the algorithm’s work on us. The preferences are prefigured in the abstraction of the school as a system and the kinds of people it is to produce.

The paradox of the systems' principles in the assessments is that the search for harmony and consensus morph into cultural practices of normalcy and pathology. Comparing occurs through the universal norms and distinctions embodied in the standards of benchmark and "the empirical evidence". Differences and divisions are not only about nations. The differences are about societies and people. The divisions from the norms formed as the populations dangerous to the system's models.

## Notes

- 1 An earlier version of this appeared in Popkewitz, Thomas S. (2017), "The promise of empirical evidence and benchmarks: The Lorelei's whispers.", in Tapio Salonen and Hans Lindquist (Eds.), *Knowledge for Change Lecture Series 2*, Malmö: Malmö University. Online publication available at <http://blogg.mah.se/knowledgeforchange/the-promise-of-empirical-evidence-and-benchmarks-the-loreleis-whispers/>.
- 2 This essay brings together different research projects related to a history of present social science/ educational reform-oriented research listed below. This includes Thomas Popkewitz and Jingying Feng work in a Swedish Science Council research project with Sverker Lindblad of the University of Gothenburg and Daniel Pettersson of the University of Gävle related to the sociology of science (International Comparisons and Re-modelling of Welfare State Education), and a book tentatively entitled "The Impracticality of Practice Research: Strategies of Change that Conserve" by Thomas Popkewitz, to be published in Chinese by Beijing Academic Press. It also includes Lei Zheng's dissertation project on historicizing the discourse about STEM crisis in education reforms.

## Notes on Contributors

*Thomas S. Popkewitz*, Professor, Department of Curriculum and Instruction, University of Wisconsin-Madison. He is interested in "how we think", focusing on the historical conditions that order the way education is thought of, acted on and talked about in current school reforms, research and teacher education.

*Jingying Feng*, Ph.D student in the doctoral program of Curriculum Studies & Global Studies at University of Wisconsin-Madison. She is currently participating as a researcher in the program "International Comparisons and Re-modelling of Welfare State Education".

*Lei Zheng*, Ph.D candidate in the doctoral program of Curriculum Studies & Global Studies at University of Wisconsin-Madison. She is currently writing her dissertation project on historicizing the discourse about STEM crisis in education reforms.

## References

- Chen, L., Dorn, E., Krawitz, M., Lim, C., & Mourshed, M. (2017). *Drivers of student performance: Asia insights*. McKinsey & Company. Retrieved from <https://www.mckinsey.com/industries/social-sector/our-insights/drivers-of-student-performance-asia-insights>.

- Cohen-Cole, J. (2014). *The open mind: Cold war politics and the sciences of human nature*. Chicago: University of Chicago Press.
- Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research*. Cambridge, UK: Cambridge University Press.
- Dewey, J. (1967–1990). Christianity and democracy. In J. A. Boydston (Series Ed.) *The collected works of John Dewey, 1882–1953*, L. Hickman (Ed.) *The collected works of John Dewey, 1882–1953 (electronic edition)*. Retrieved from <http://www.nlx.com/collections/133>.
- Ministry of Education of the People's Republic of China. (2015). *Unmasking PISA of Chinese characteristics*. Retrieved from [http://www.moe.gov.cn/jyb\\_xwfb/xw\\_fbh/moe\\_2069/xwfbh\\_2015n/xwfb\\_150415/150415\\_mtbd/201504/t20150416\\_187197.html](http://www.moe.gov.cn/jyb_xwfb/xw_fbh/moe_2069/xwfbh_2015n/xwfb_150415/150415_mtbd/201504/t20150416_187197.html). (in Chinese)
- Mourshed, M., Chijioko, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. Chicago, IL: McKinsey & Company.
- Mourshed, M., Farrell, D., & Barton, D. (2013). *Education to employment: Designing a system that works*. Chicago: McKinsey & Company.
- OECD. (2017). *PISA 2015 results (Volume III): Students' well-being*. Paris: OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/9789264273856-en>.
- Pont, B., Donaldson, G., Elmore, R., & Kools, M. (2014). *The OECD-Sweden education policy review: Main issues and next steps*. Paris: OECD.
- Popkewitz, T. (Ed.). (2005). *Inventing the modern self and John Dewey: Modernities and the traveling of pragmatism in education*. New York, NY: Palgrave Macmillan.
- Popkewitz, T. (2008). *Cosmopolitanism and the age of school reform: Science, education, and making society by making the child*. New York, NY: Routledge.
- Popkewitz, T. (2010). The university as prophet, science as its messenger, and democracy as its revelation: John Dewey, University of Chicago President William Rainey Harper, and Colonel Francis Parker. In T. Schlag, D. Tröhler, & F. Osterwalder (Eds.), *Pragmatism and modernization* (pp. 99–122). Rotterdam: Sense Publishers.
- Wang, H. (2006). Discursive community and the genealogy of scientific categories. In M. Y. Dong & J. Goldstein (Eds.), *Everyday modernity in China (Studies in Modernity and National Identity; A China Program Book)* (pp. 80–117). Seattle: University of Washington Press.
- Wang, H. (2008). *Rise of modern thoughts in China*. Beijing: Joint Publishing Company. (in Chinese)