Excellence in Higher Education

Excellence in Higher Education 7 (2016): 7-26

Konsorsium Perguruan Tinggi Indonesia-Pittsburgh

An International Comparison of the Educational Research Quality Standards/Indicators Scheme

Shen-Keng Yang[a,*]

[a]*National Taiwan Normal University, Taiwan*

Abstract

This paper investigates the establishment of a quality research scheme that draws from optimal educational research from the European Union, Organisation for Economic Co-operation and Development (OECD), Australia, France, Germany, the United Kingdom, and the United States. Based on the results of one of the sub-projects of the Integrated Joint Research Project financed by the Science Council in Taiwan, this paper focuses on the process and construction of an Educational Research Quality Standards/Indicators Scheme. The implications of the results of this comparative study will help improve education research generally and will provide suggestions for future use in assessing the inputs and outputs of research projects funded by governmental and non-governmental agencies in Taiwan and in international contexts.

Abstrak

Makalah ini menelaah rancangan skema kualitas riset yang merujuk pada riset kependidikan di Uni Eropa, OECD (Organisasi Kerjasama Ekonomi dan Pembangunan), Australia, Perancis, Jerman, Inggris dan Amerika Serikat. Berdasarkan temuan salah satu subdivisi Proyek Riset Bersama (Integrated Joint Research Project) yang dibiayai oleh Dewan Sains di Taiwan, makalah ini difokuskan pada proses dan pengembangan Skema Indikator/Standar Kualitas Riset Kependidikan. Implikasi temuan studi komparatif ini akan membantu meningkatkan kualitas riset pendidikan secara umum, sekaligus menawarkan rekomendasi terkait dengan evaluasi input dan hasil riset yang dibiayai oleh unit pemerintah maupun non-pemerintah di Taiwan dan di forum internasional.

Keywords: Education Research, Education Research Policy, Standards, Indicators, Education Research Quality, Taiwan

Introduction

Under the impacts of the rapid progress of science and technology and the quick pace of globalization, many countries worldwide have adopted strategic plans for social, cultural and economic development to outstrip rivals in the fiercely competitive international arena. Education has been thought to play a key role in fostering social, cultural and economic development. Being aware of the importance of education to national development, many countries take education reform as their administrative priority.

Education reform should be based on good education research in order to have sound judgment about the strengths and weaknesses of the educational policies and practices in a country. Unfortunately, the quality of education research has been criticized for its sloppiness and uselessness in many countries as well as in international organizations. In OECD's (1995) report *Educational Research and*

*Corresponding author. Address: National Taiwan Normal University, Taiwan. Email: skyang@ntnu.edu.tw.

ISSN 2153-9669 (print) 2153-9677 (online) | doi: 10.5195/ehe.2016.150 | **http://ehe.pitt.edu**

Development: Trends, Issues, and Challenges, the production of educational research was criticized as very weak and thus leading to a patchwork of educational knowledge. The European Commission (2008) has also criticized educational research in European countries as locked into national intellectual resources and lack of international visibility (European Commission 2012, p. 29). In the 2012 report *Global Europe 2050*, most of the research governance in Europe is judged as fragmented within national borders even more than in the 2010s.

Critique leveled against the fragmentation, poor quality and the gap between research and policy/practice in education research has been also raised in many countries. In the UK, Hargreaves (1996) criticized educational research in his *Annual Lecture to the Teacher Training Agency* as second-rate, without contribution to fundamental theory, irrelevant to practice, and uncoordinated with any preceding or follow-up research. In response to the criticism and caring about the future of education research, Her Majesty's Chief Inspector of Schools, Chris Woodhead, commissioned J. Tooley and D. Darby (1998) through the Office for Standards in Education (OFSTED) to carry out a study on the quality of educational research. The results of the study find educational research full of partisanship, logically incoherent, without connection to practice, and sloppy in methodology. In the United States, as Atkinson and Jackson (1992, p. 19) indicate, research in education is likely to be dismissed as trivial or irrelevant to practice. In Germany, a series of workshops from June 2004 to January 2006 was launched by Hochschulrektorenkonferenz (HRK) to examine the lack of coordination between educational research institutes and the low quality of educational studies. In France, Antoine Prost was commissioned to write a report on the strategic program for educational research and the report concluded that educational research in France was rather weak, fragmented and too dispersed to be cumulative.

Being aware of the importance of high-quality educational research to the improvement of educational practice and of the poor quality of recent educational research, international organizations and many countries have established special institutes responsible for monitoring educational research and its use in educational practice. For the purpose of controlling the quality of educational research, international organizations and many countries have founded special responsible agencies in charge of educational studies (e.g., the Centre for Educational Research and Innovation of OECD, the Office of Educational Research and Improvement in the United States). These organizations have developed a series of quality indicators of various types and for various levels of educational research. This paper addresses itself to the comparison of these various kinds of quality indicators. First, critique of educational research quality will be analyzed, which will be followed by the explication of the meanings of quality, excellence and standards. The paper then moves to the exploration of the founding of special responsible agencies for monitoring educational research quality by international organizations as well as by governmental and non-governmental research units. The quality indicators/standards developed by international organizations, university departments and independent research units will be carefully compared, with a view to constructing a sound "Educational Research Quality/Indicator Scheme" for assessing educational research in the future.

Criticism of Educational Research Quality

The awful reputation of educational research has a long history. In the inaugural issue of *Educational Review*, Josiah Royce (1891, pp. 23-24) published an article entitled, "Is there a science of education?" In the essay, Royce was of the opinion that "there was no universally valid science of pedagogy... capable of... complete formulation and... direct application to individual pupils and teachers" (p. 16). In a word, Royce criticized the studies in education of that time as unsystematic, without the possibility of being put into practice.

From an historical perspective, Lagemann (2000, p. ix) observes also that, though education emerged as a subject of university studies, there were neither focus and unified methods of investigation, nor any coherent, strong self-regulating professional communities.

Though great progress in educational research was made in the later half of the last century, in 1960 Kerlinger still leveled his sharp criticism toward the educational research community. He criticized education research as fraught with mythology, that is, "a body of legends and beliefs purporting to be the rationale, purpose and methods of educational research" (1960, p. 149). Three research myths pertained to "methods," "practicality," and "statistics." The methods mythos is centered on the naive misconception that research design is synonymous with research methodology. For Kerlinger, more attention should be put on the investigated questions, rather than methods. The "practicality" mythos referred to the preoccupation with usefulness when designing, conducting, or evaluating research. Practical research is very important for educational practitioners and policy-makers. However, basic research is more promising than applied research as a means for understanding educational phenomena. Concerning the statistics myth, Kerlinger meant two problems: (a) a fundamental disregard for statistics as an informational and methodological tool, with the result being a multifarious series of numerical abstractions that have little or nothing to do with educational reality; and (b) a failure to understand that research design and statistical analysis are intimately related, with the result of causing researchers to settle for sophisticated data analytic methods and/or doing a poor job of interpreting the results of the methods they use.

The criticisms of educational research quality have recurred throughout the twentieth century, growing even sharper today. In the United States, the Congress founded the National Institute of Education in 1992 because they believed that the research programs of the Office of Education were mediocre and useless. In 1971, the Presidents' Commission on School Finance commissioned the RAND Corporation to review research on what is known about what works in education (Averch et al. 1972). The result of the review found that the body of educational research available at that time was very unsatisfactory as compared with studies in numerous other fields. In 1972, the RAND Corporation commissioned Averch and colleagues (1972) to conduct research on the effectiveness of schooling. They found that there is no consistent and unambiguous evidence to support enhancement of effect and efficiency of students' learning. The Report of the National Science Foundation (NSF) (1999), "*Improving Student Learning: A Strategic Plan for Education Research and Its Utilization*," has a similar conclusion as the Rand Report: educational research is not rigorous enough for the betterment of educational practice. Further, the National Research Council (NRC) (1999) also criticized educational policy-making in the US as based on personal experience and ideology, with the consequence of inadequacy and uselessness in improving educational practice. In 2002, the NRC released "Scientific Research in Education." In addition to explicating the nature of scientific educational research, the Report aimed at guiding efforts to improve educational research quality. According to the Report, the findings from education studies "are of low quality and are endlessly contested the result of which is that no consensus emerges about anything" (NRC 2002, p. 28). In other words, the criticism leveled at educational research is that its results cannot contribute to the growth of knowledge and cannot provide any useful information for improving policy-making and educational practice.

Parallel to the trends of critique of educational research quality in the US, there arose also an indictment against the current situation of studies in education in the UK. The critique of educational research in recent years began when David Hargreaves (1996, p. 7) delivered his *Annual Lecture to the Teacher Training Agency* with a sharp criticism of the educational research community. The criticisms leveled at educational research offered by Hargreaves included the following points:

1. Educational research is second-rate, without any serious contribution to fundamental theory or knowledge;
2. It has no relevance to educational practice;
3. It is uncoordinated with any preceding or follow-up research;
4. Mediocre research clutters up academic journals that virtually nobody reads.

In light of three of Hargreave's significant criticisms, they built criteria for "good practice" in educational research in three categories: (a) whether educational research makes a serious contribution to a fundamental theory of knowledge, (b) whether educational research is irrelevant to practice, and (c) whether educational research is uncoordinated with any preceding or follow-up research. In each category, some questions were proposed in terms of focus, conduct and presentation of research. For example, the "good practice" criteria questions concerning the conduct of the research within the category of "making a serious contribution to fundamental knowledge" include the following, which should ideally merit affirmative answers:

- Does the research involve triangulation to establish the trustworthiness of its findings?
- Does the research avoid a sampling bias?
- Does the research use primary sources in the literature review?
- Does the research avoid partisanship in the way the research is carried out and in the interpretation of data?
- Do the conclusions follow from the evidence presented?

In the other two categories of "good practice" in educational research (i.e., "relevant to practice" and "coordination with preceding and follow-up research"), some questions were also proposed to check whether the research was good enough or not.

With this list of questions, Tooley and Darby selected 264 articles from four representative British educational journals to check if they met the requirement of good educational research. The results were very unsatisfactory. As Tooley (2001, p. 138) points out, the study reveals that there is a large amount of second-rate academic research in terms of partisanship, methodology, and the argumentation of non-empirical research. Many selected articles in this study did not present factual details about sample size and sampling method. Furthermore, most articles have shown themselves as having the severe weakness of educational research that is conducted as if carried out in a vacuum, often apparently unheeded and unnoticed by anyone else in their field. In a word, most articles under the study are full of partisanship, logically incoherent, methodologically sloppy, and uncontexualized in their research.

Furthermore, the Department of Education and Employment commissioned J. Hillage, R. Person, A. Anderson and P. Tamkin (1998) to produce a report titled *Excellence in Research on Schools*. They found also that almost all those involved with educational research, including practitioners, decision-makers, researchers, sponsors and beneficiaries, were very dissatisfied with its quality and usefulness.

In response to the criticisms, a series of measures for improvement of educational research have been launched by the British Educational Research Association, the Academy of Learned Societies in the Social Sciences and other related organizations. The situations of educational research in the UK have taken a turn for the better in the last ten years. However, due to political change and economic recession, the public status of educational research is still fraught with uncertainties and tensions (Oancea 2011, p. 6).

The situation of educational research in France is much worse even than in the UK. In 2001, Minister of Education Jack Lang commissioned Antoine Prost (2001) to write a report on the strategic program of educational research. The result revealed that educational research in France was very weak, fragmented and too dispersed to demonstrate cumulative growth in knowledge. Though measures have been taken to amend the drawbacks of educational research, the situation has not changed much. In the last ten years more than 30 educational research centers have emerged and there are about 600 to 800 researchers who have devoted themselves to educational research. There is still no apparent improvement in educational studies. Further, "science de l'éducation" was created by ministry decree as an institutional discipline in universities in order to provide an academic framework for teacher training. Nevertheless, the term "science de l'éducation" has no clear epistemological consensus: is it defined by object? by methods? by values? or by common references? There is no common agreement on these issues in French educational science. Even worse, on an academic level, studies carried out in the circle of educational science are often ranked lower than those conducted in other disciplines (e.g., sociology, psychology, or cognitive science). The politicians and policy-makers are often suspicious of educational research, and sometimes seek legitimate policies for research that appears more scientific, such as cognitive psychology or neuroscience. Educational research is actually viewed not as a strong "scientific" field, but rather an ideological one (Rey 2011, pp. 28-29).

In Germany, though scientific educational research can trace its origins to the eighteenth century when the first academic journal (perhaps also the first one in the world), "*Magazin für Schulen und die Erziehung überhaupt*" was published in 1766, its shortcomings and irrelevance to policy and practice are still raised in critiques. In the inaugural issue of the first journal, it announced that *Erziehungswissenschaft* (the science of education—a term that seems to have appeared for the first time in academic history) was the favorite term of the journal. The main purpose of the journal was to publish articles concerning the result of scientific studies. Through scientific studies the principles and rules of educational process could be established. Observing these principles and rules in their teaching, teachers could avoid appealing to pedantry's authority, to bias, to ideology or, even worse, to superstition in managing their school activities. Nevertheless, educational science or educational research today in Germany cannot escape critique.

From June 2004 to January 2006, Hochschulrektorenkonferenz (HRK, Conference of University Presidents), in cooperation with universities, launched a series of workshops to examine the weaknesses of teaching and research and the lack of coordination of different units of educational research with a view to improving the quality of research and teaching. Stefan Hornbostel (2004, p. 78) indicated also that educational science in Germany, owing to its traditional affiliation with humanistic science (*Geisteswissenschaft*) and cross-disciplinary characteristics, could not be completely internationalized. In order to rectify this weakness, the Federal Ministry of Education and 16 State Ministries of Education signed an agreement (*Hochschulpakt 2020*) with the aim of improving university research facilities and providing extra-expenditures for farsighted research projects to enhance the international visibility of German research products.

Not only various countries, but also international organizations, have raised critiques leveled at educational research quality. The Organization for Economic Co-operation and Development (OECD 2007, p. 16), after country reviews of educational R&D, confirmed the following features as commonly characterizing OECD systems:

- Low levels of investment in educational research;
- Generally low levels of research capacity, especially in quantitative research; and
- Weak links between research, policy, and innovation.

Similar critiques of educational research have also been raised by the European Commission (2012, p. 80) in the *Global Europe 2050*:

> the innovation system thinking and coordination capabilities at the EU level fail to emerge, leaving the whole of Europe in an unfavorable competitive position as compared to other regions of the world, and especially to emergent economies.

Measures have been taken by the EU to improve these deficiencies of research, especially with the establishment of the European Research Area (ERA) for the purpose of integrating research conducted in European countries. However, public sector agencies of various countries collect and diffuse data in a fragmented manner and with varying pricing practices for public sector information in quite different ways. Even worse, duplications and the fragmentation of educational studies are still found in many countries, with the consequence of wasting limited resources. Furthermore, the gaps between research and practice are still large, waiting for amendment.

Quality, Standards, and Indicators

The foregoing analysis of the critiques and debates on educational research quality shows that all those involved in the discussions have no commonly agreed meanings of the terms *quality*, *standards*, and *indicators*. It is therefore imperative to clarify the meanings of these terms before going on to investigate the quality standards of good educational research.

The term quality in Greek is *ποιὸν* or *ποιότης* can trace its origin to Democritus (around 460-370 B.C.), who distinguished between primary qualities based on the shape and characteristics of the *atomon* and secondary or derived qualities (like sweet, bitter, warm, cold, etc.), which are conventional and essentially subjective and passive. It was Plato who first used the abstract form of the term *ποιότης* to differentiate it from substance (Plato, Timaieus, 49a-50a). The quality *ποιὸν* in Aristotle's philosophy is one of the ten categories used to describe *being* (Aristotle, De Cat, 1b-2a). The Greek words for quality (*ποιὸν* and *ποιότης*) are predicative to describe an object.

How does one determine if the quality is good or bad, positive or negative? There must be some standards or criteria for reference of sound judgment. Etymologically, the term *standard* has no correspondent origin in either Greek or Latin. According to Johannes Bilstein (2009), standard can trace its origin to Old French *standard*, which means "a place for lining up or erection." In military terminology, it means a place where a flag is erected as a guideline for gathering the soldiers. From Old French, *standōrd* transforms into the English standard, which often means an optical and spatial point of reference or guideline. The point or guideline can help people to take common action together. The traditional French and English term for standard transformed in the nineteenth century to the German *standard* as a merchandise indicator, which means a prototype or model guaranteeing the quality of the goods.

From an educational perspective, quality and its standards are much more complex and multifarious. Quality may take meanings depending upon: (a) the understandings of various interests of different constituencies or stakeholders in a given institute; (b) its reference: inputs, outputs, mission, objectives, etc.; (c) the attributes or characteristics of the academic world that are worth rating; and (d) the historical period in the development of higher education and research organization.

UNESCO has thus proposed a spectrum of definitions of academic quality as follows:

1. Quality as excellence: a traditional, elitist academic view, according to which only the best standards of excellence are understood as revealing true academic quality;
2. Quality as fitness *for* purpose: a concept that stresses the need to meet or conform to generally accepted standards as those defined by an accreditation or assurance body, the focus being on efficiency of the process at work in institutions or programs in fulfilling the stated, given objectives and mission;
3. Quality as fitness *of* purpose: a concept that focuses on the defined objects and mission of a given organization; and
4. Quality as enhancement or improvement: focusing on the continuous search for permanence, stressing the responsibility of higher education institutes and research organizations to make the best use of academic freedom and autonomy. (Vlãsceanu et al. 2007, pp. 70-73)

Harvey and Green (1993) also proposed five categories of quality as follows:

1. The exceptional view of quality as *excellence* unattainable by most;
2. Quality as *perfection* views quality as a consistent or flawless outcome;
3. Quality as fitness for purpose sees quality in terms of fulfilling a customer's requirement, needs or desires;
4. Quality as value for money sees quality in terms of return on investment; and
5. Quality as transformation denotes that it can change one situation to another. In educational terms, it refers to the enhancement and empowerment of students, teaching and research staff of an institute as well as the growth of knowledge.

Not only are there debates on the meanings and definitions of research quality, but also there are various discussions on the issues of the approach and the methodology of defining or deciding the quality of educational research. Traditionally, peer review, bibliometrics, impact factor of publications and articles published in SCI or SSCI journals are used to assess research results, research proposals with the purposes of recruiting new researchers, promoting the academic status of researchers or allocating research funds.

However, all the mentioned approaches and methodologies are not without questions. Most articles published in SCI or SSCI are confined to the English language. Only a few articles are published in German or another European language. It is thus unfair to those competent researchers who have an acquaintance with languages other than English. Furthermore, in some disciplines (such as philosophy of education, history of education, etc.), publications in book form or book chapter are much more important than journal articles.

Peer review, as a commonly used method to control research quality, is also full of problems. The first issue confronting peer review is how to select competent reviewers. Even if the chosen reviewers are very famous for research in their reviewed disciplines, they cannot be completely free from personal bias and ideologies. Panel review is also problematic or even worse. Since the constituent members are selected from different institutes, partisanship and tribalism cannot be completely avoided in the reviewing process. In order to protect from possible bias and partisanship, setting up objective standards for reviewer's reference is the prerequisite to guaranteeing the quality of educational research under review. However, the question of who or which organizations are authorized to establish educational research quality is also a big problem. These issues are discussed in the next section.

Institutes Responsible for Quality Assurance in Educational Research

In light of the important role of educational research in determining the improvement of educational policies and practice, thus leading to social, cultural and economic progress, many international agencies and governments in various countries have established institutes to monitor the quality of educational research. Especially confronting the emerging new development issues in the twenty-first century, the research institutes must undertake the soaring responsibility for finding possible solutions to these problems. A good quality of educational research is the key to flourishing social, cultural and economic development. The research institutes at international, national, as well as private levels thus play key roles in monitoring research and knowledge systems for complex educational and social activities.

OECD countries, in consideration of the patchy works and lack of coordination in educational research and policy-making in the member countries, established the Centre for Educational Research and Innovation (CERI) with the main mission of opening up new fields for exploration and combining rigorous analysis with conceptual innovation. The Centre's extensive work covers learning at every age, from birth to old age, and also goes beyond the formal education system. CERI has a particular concern with emerging trends and issues in education resulting from the rapid pace of globalization and the quick progress of science and technology. It puts specific emphasis on accumulating statistical evidence for the value of its research work.

To keep a high level of educational research and effectiveness of knowledge utilization, CERI draws on extensive expertise and methodologies from across the OECD area. Its work is designed to satisfy the needs of all stakeholders, especially policy-makers, research communities and leaders. Currently, five main projects, including the topics of education system governance, innovation strategy, social progress, innovative learning environment, and innovative teaching and effective learning are underway to achieve the main goals of CERI: to generate forward-looking research analyses and syntheses, to identify and stimulate educational innovation, and to promote international exchange of knowledge and experience.

For enhancing the international competitiveness of European countries and coordination of research, innovation and policy of the member countries, the European Commission has established the Joint Research Centre. It comprises seven research institutes located in five EU member states (Belgium, Germany, Italy, the Netherlands, and Spain). Its mission is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. The Centre is governed by a Board of Governors, comprised of high-level representatives from EU member states, candidates and associated countries. The main mission of the Board is to advise on strategy, work program, budget and high-level appointments. Under the important consultation Seventh Framework Programme (FP7), a series of standards for assessing educational research qualities have developed. The standards will be discussed in the next section.

Not only international organizations, but also many countries have established institutes responsible for monitoring educational research quality. In the US, the Office of Educational Research and Improvement affiliated with the U.S. Department of Education is the federal government's lead agency for educational research and development. Its main goals are to promote quality and equity in education. OERI's immediate predecessor was the National Institute of Education (NIE), created in 1971 with the mission to provide "leadership in the conduct and support of scientific inquiry into the education process" and to build "an effective educational research and development system." Since its inception, NIE has scarcely been free from political interventions. NIE was always vulnerable to charges that its research programs were influenced by the political and ideological concerns of the administration, congressional sponsors, and agency managers (Atkinson and Jackson 1922, pp. 57-58). Owing to its small scale and limited budget, NIE became one of the institutes belonging to the "holding company," the Office of Educational Research and Improvement, created in 1979.

The occasion of establishing a national educational research quality monitoring institute like OERI can be traced its origin: *An Act to Establish a Department of Education* in 1867, which stipulated the establishment of the Office of Education (USOE). Its mission was to collect such statistics and facts as to show the condition and progress of education, to diffuse information about the organization and management of schools and school systems, teaching

methods, and to aid the people of the United States in the establishment and maintenance of efficient school systems.

However, USOE's research activities were primarily restricted to the routine collection and dissemination of statistics in its first nine decades. The federal investment in education research was minimal. Only with the stipulation of the 1954 *Cooperative Research Act* was USOE authorized to provide funds for field-initiated research, primarily at universities, much as other federal agencies were doing for research in the natural sciences. Under the Act, individual projects were funded through proposals initiated from the field, with little opportunity for federal officials to shape a national research agenda. Systematic research for improving educational policies and practice was still lacking.

In 1979, the OERI was established when the Office of Education was replaced with a Department of Education. It was originally seen as a "holding company" for NIE, the National Center for Education Statistics (NCES), Library Programs, and some other discretionary and dissemination activities. OERI was to provide some overall guidance and coordination, but to allow the main entity to operate semi-autonomously. OERI was restructured in 1985 into five offices.

Nowadays, OERI is composed of six offices: the Office of the Assistant Secretary, the Office of Research, Programs for the Improvement of Practice (PIP), the National Center for Education Statistics (NCES), the Fund for the Improvement and Reform of Schools and Teaching (FIRST), and Library Programs. In addition to achieving the goals of the Congressional mandate to promote the quality and equity of American education, OERI has the mission to support educational research of the highest quality and to strengthen the educational research and development system.

In addition to national institutes of educational research quality monitoring like OERI, NIE, etc., there are also many private or independent organizations of educational research and development. One of the oldest and most important research institutes is the RAND Corporation. The primary idea of creating an independent organization for connecting military planning with research and development decisions, like RAND, can trace its origin to some farsighted persons working at military units and private organizations during the close of World War II. World War II revealed the importance of technology research and development for success on the battlefield. In order to assure complete and permanent peace, the forward-looking individuals, General H.H., Hap Arnold, Edward Bowles, Donald Douglas, Arthur Raymond, and Franklin Collbohm met to set up Project RAND under a special contract to the Douglas Aircraft Company. The project stressed the importance of coordination, balance, and cross-fertilization of efforts from the laboratories of industry, the seminars of universities, and the offices of administration.

On 14 May 1948, RAND was incorporated as a nonprofit corporation under the laws of the State of California. The goal of RAND was "to further and promote scientific, educational, and charitable purposes, all for public welfare and security of the United States of America." For securing the quality of RAND's publications and productions, the Corporation decided to codify in writing the quality standards for all RAND research. The written standards were put on RAND's web page in 1997 for public discussion. The newly revised version in 1999—*RAND Standards for High-Quality Research and Analysis*—will be discussed in the next section.

In the UK, an independent educational research unit was created in 1946 in the name of the National Foundation for Education Research following *The Education Act of 1944*, which introduced free secondary education for all and included powers for the Ministry of Education and local education authorities to fund research on educational provision. By 1949, the Foundation was a self-governing body supported by all local education authorities, universities and national teachers associations, which became corporate members. In 1967, NFER became an incorporated, charitable body. Nowadays, it is an independent charity and active fundraiser working to provide evidence that improves education and learning and, as a result, the lives of learners. Through providing unbiased expert consultation, the efforts to connect to influential governmental departments, stakeholders, policy-makers, practitioners, and its emphasis on accountability in educational research and planning, the NFER plays a key role in monitoring the quality of educational research and educational activities.

In Germany, there are also various institutes for monitoring the quality of educational research and educational practice at federal, state and private levels. One of the bigger institutes is Deutsche Forschungsgemeinschaft (DFG 2011). It is an independent private legal corporation financed mainly by the federal and state governments. In addition to supporting individual researchers and strengthening international exchanges of research information, the DFG also emphasizes quality control through rigorous review of research proposals.

Deutsches Institut für Pädagogische Forschung is a non-university scientific institute that offers infrastructural service for research, praxis, administration, and policy in education systems. It provides the public with national information on education research, on the one hand, and carries out research on and evaluation of the education system, on the other.

The Max Planck Institut für Bildungsforschung (MPIB) is a multidisciplinary establishment with about 80 research facilities financed by Max-Planck-Gesellschaft Zur Förderung der Wissenschaften e.v., the core support for which is provided by federal and state governments. The inquiries of MPIB concentrate on the evo-

lutionary, social, historical, and institutional contexts of human development, as well as examining them from life span and life course perspectives. Quality control is carried out in cooperation with foreign experts in reviewing the processes and outcomes of the conduct of research projects.

The Forschungsinstitut für Bildungs und Sozialökonomie is a private and independent research and consultation institute founded by Dieter Drohmen in 1993. The institute emphasizes research on all economic aspects of education, social problems, workforce, and demographic changes.

Standards of Educational Research Quality from an International Perspective

For purposes of monitoring educational research quality, international organizations and various countries have developed standards of research quality to assess research proposals, processes and production at different levels. In seeking to correct the weakness and patchwork of educational research in its member states, OECD invited expert researchers and statisticians to have a consultant meeting in Frascati, Italy in 1963 to jointly write the *Proposed Standards Practice for Surveys of Research and Experimental Development*, better known as the *Frascati Manual* for short. The *Frascati Manual* serves as the standard criteria based on which scientific research systems, research policies, research personnel, financial allocation and research productions in various countries are assessed. Based on the *Frascati Manual*, UNESCO formed a Group of National Experts on Sciences and Technology Indicators to develop a series of manuals for research methods, research quality, and research policy as a reference for researchers in different countries.

With a view to building a European Research Area, the European Union has launched the Seventh Framework Programme (FP7) to strengthen European scientific and technological research. Under the FP7, the European Educational Research Quality Indicators (EERQI) Project was initiated. A series of indicators for assessing educational research was proposed for the discussions in a two-day workshop held at Leuven University. The series is composed of extrinsic and intrinsic indicators. The extrinsic indicators refer to citation index, impact factor, etc. The intrinsic factors are the essential characteristics of the research, which consist of five categories: rigor, originality, significance, integrity, and style. After a series of discussions and consultations, integrity and style were revoked because they are not transparent, nor were they considered valid criteria for assessing research. Sixteen concrete indicators within the categories of rigor, originality and significance were proposed to evaluate not only empirical, but also philosophical and historical studies of education.

In the United States, the NRC of the National Academy of Sciences commissioned Shavelson and Towne (2002) to examine and clarify the nature of scientific inquiry in education and how the federal government can best foster and support it. After careful study, six fundamental principles underlying scientific research in education were proposed:

1. Pose significant questions that can be investigated empirically,
2. Link research to relevant theory,
3. Use methods that permit direct investigation of the question,
4. Provide a coherent and explicit chain of reasoning,
5. Replicate and generalize across studies, and
6. Disclose research to encourage professional scrutiny and critique.

RAND, as an independent, nonprofit, charitable corporation, commits itself to high-quality, objective research and analysis on issues at the top of the national and international policy agenda. To assure the high quality of research and analysis, RAND has made public its standards for high-quality research and analysis since 1997. The standards are important tools, not only for those who are involved in Rand research production and research utilization, but also generally for those who have a common interest in educational research. The standards of RAND's high-quality research and analysis run as follows:

1. The problem should be well formulated, and the purpose of the study should be clear,
2. The study approach should be well-designed and executed,
3. The study should demonstrate understanding of related studies,
4. The data and information should be the best available,
5. Assumptions should be explicit and justified,
6. The findings should advance knowledge and bear on important policy issues,
7. The implications and recommendations should be logical, warranted by findings, and explained thoroughly, with appropriate caveats,
8. The documentation should be accurate, understandable, clearly structured, and temperate in tone,
9. The study should be compelling, useful, and relevant to stakeholders and decision-makers, and
10. The study should be objective, independent, and balanced. (RAND Foundation 2011, p. 19)

RAND assures that its production has not only short-term value in response to the immediate policy concern, but also enduring

value in furthering the growth of scientific knowledge. This is the reason why RAND places its research products in the public domain. RAND's standards of high-quality research and analysis may be universally very useful for those who engage in promoting the progress of educational knowledge through objective research.

For the purpose of assisting researchers in the preparation of manuscripts that report work in empirical education research, editors and reviewers in the consideration of these manuscripts for publication, and readers in learning from and building upon such publications, the American Educational Research Association (AERA 2006) published *Standards for Reporting on Empirical Social Science Research in AERA Publications*. The reporting standards are divided into eight general areas: problem formation; design and logic of study; sources of evidence; measurement and classification; analysis and interpretation; generalization; ethics in reporting; and title, abstract, and headings. Each area starts with a general discussion about the research, followed by specific numbered standards that pertain to that domain. In total, 43 standards pertaining to the above eight areas are proposed to provide a framework for what a report of empirical work ordinarily should address.

The AERA (2009) published *Standards for Reporting on Humanities-Oriented Research in AERA publications* to complement the *Social Science Standards*. The document, *Humanities-Oriented Standards*" has two sections. The first section describes humanities-oriented research in terms of its primary methods, purposes and content, as well as its inherent controversies. In the second section of the report, seven standards, each with a series of sub-standards, for a total of 43 sub-standards that explicate and elaborate the major standards are set forth: (1) significance, (2) methods, (3) conceptualization, (4) substantiation, (5) coherence, (6) quality of communication, and (7) ethics. The purpose of the proposed series of standards is to assist those who engage in humanities-oriented education research, including reviewers, editors, readers and researchers, to have a whole picture of humanities-oriented educational studies.

In light of the importance of ethics in the production and dissemination of education research, the AERA (2011) published in 2011 *The Code of Ethics of the American Educational Research Association*, which articulates a common set of values upon which education researchers build their professional and scientific work. The *Code* consists of five principles: (1) professional competence; (2) integrity; (3) professional, scientific, and scholarly responsibility; (4) respect for people's rights, dignity, and diversity; and (5) social responsibility.

With these principles reflecting the highest ideals of professional conduct, the AREA expounds 22 standards as rules for ethical conduct by education researchers. The ethical standards are set forth as follows: (1) scientific, scholarly, and professional standards; (2) competence; (3) use and misuse of expertise; (4) fabrication, falsification, and plagiarism; (5) avoiding harm; (6) nondiscrimination; (7) non-exploitation; (8) harassment; (9) employment decisions; (10) conflicts of interest; (11) public communications; (12) confidentiality; (13) informed consent; (14) research planning, implementation, and dissemination; (15) authorship credit; (16) publication process; (17) responsibilities of reviewers; (18) teaching, training, and administering education programs; (19) mentoring; (20) supervision; (21) contractual and consulting services; and (22) adherence to the ethical standards of the American Educational Research Association. Most of the ethical standards are written broadly in order to apply to educational researchers in varied roles, and the application of an ethical standard may vary depending upon the context.

The foregoing standards of scientific and humanities-oriented research, as well as standards of professional ethical conduct of researchers, are proposed to provide guidelines mainly for those engaged in the reporting, reviewing, editing and disseminating of their research in AERA publications. Nevertheless, in light of their breadth in the spectrum of application and their profundity in explication all the mentioned standards, both scientific and ethical, are applicable to other forms of research documents worldwide. The standards deserve careful deliberation by those who engage in monitoring educational research quality at various levels, national or international, governmental or non-governmental.

In the UK, an important component of reforms of public sector financing was the introduction of devolved management accompanied by increased selectivity and concentration, and a move from negotiated budgets to "performance budgeting" throughout the 1980s and 1990s. Performance assessment of higher education has thus become a very important way for those deserved institutes to earn necessary financial resources in a fair way. Public evaluations based on externally decided benchmarks and indicators have to be institutionalized to guarantee the objectivity of the assessment. The Research Assessment Exercise, first introduced in 1986, played an important role in ranking higher education institutes in the last three decades.

The UK Research Assessment Exercise was carried out by assessment panels that were constituted on the basis of nominations from learned societies in relevant fields, but additionally with representatives from the "users'" communities for research in those fields. Their size reflects the likely scale of the work to be submitted in any particular field. Education was one of the largest fields, joined by psychology and sport science.

The higher education institutions were asked to provide the following data as a basis for the assessment: information about the numbers of submitted staff under different categories; quantitative data related to research students and studentships; quantitative data

about research income and its sources; and a descriptive account of the research environment.

A five-point scale was provided by the Higher Education Funding Councils (HEFC) for judgment criteria for each invited panel conducting an assessment:

4. Quality that is world-leading in terms of originality, significance, and rigor.
3. Quality that is internationally excellent in terms of originality, significance, and rigor.
2. Quality that is recognized internationally in terms of originality, significance, and rigor.
1. Quality that is recognized nationally in terms of originality, significance, and rigor.

Unclassified. Quality that falls below the standard of nationally recognized work.

Based on the data provided by the assessed higher education institutes, the sub-panels produced three profiles:

1. The profile of *research environment*: the profile is formed on the basis of the institute's account of its structure and environment.
2. The profile of *esteem*: the profile is formed on the basis of the institute's account of the achievement, evidence of esteem and strategy for future development.
3. The profile of *research quality*: the profile is formed on the basis of the published works submitted by the assessed institutes. The sub-panels read most of the submitted publications in order to have a fair and sound judgment.

The criteria for the assessment of the "output" employed in making the sub-panels' judgments are as follows:

1. **Originality:** a characteristic of research that is not merely a replication of other work or simply applies well-used methods to straightforward problems, but that engages with new or complex problems or debates and/or tackles existing problems in new ways.
2. **Significance:** judged in different ways according to whether the research is basic, strategic or applied. Ways of evaluating significance include judging the effects of the research on the development of the field, examining contributions to existing debates, and assessing its impact on policy and practice.
3. **Rigor:** judged in many ways, and can hopefully be associated with methodological and theoretical robustness and the use of a systematic approach. It includes traditional qualities such as reliability and validity, and also qualities such as integrity, consistency of argument, and consideration of ethical issues.

Finally, the three profiles will be combined into a single profile. The weighting for each element is determined by each sub-panel; the weighting agreed upon for the Education sub-panel is 70 percent for research outputs, 20 percent for research environment and 10 percent for indicators of esteem.

The expert panels will produce for each unit of assessment of a single numerical profile of their research quality expressed as a percentage of each of the five grades. Concerning the allocation of the financial resources based on the assessment, the four national funding councils will each make their own decisions about how to relate funding to these profiles.

The Research Assessment Exercise began in 1986, was carried out in 1989, 1992, 1996, and 2001, and the last Exercise was completed in 2008. In December 2006 the Department for Education and Employment announced that a new framework for research assessment and funding would replace the RAE after the 2008 exercise in England. The underlying policy of allocating research funding selectively on the basis of quality remained unchanged. The main intention was to make the assessment mechanism simpler and less burdensome.

The new framework for research assessment and funding is the Research Excellence Framework (REF), which was completed in 2014. The key aims of the new framework, as indicated in the Circular Letter (Number 06/2007) by the Chief Executive of HEFCE, Professor David Eastwood, are the following:

- To produce robust UK-wide indicators of research excellence for all disciplines, indicators that can be used to benchmark quality against international standards and to drive the Councils' funding for research.
- To provide a basis for distributing funding primarily by reference to research excellence, and to fund excellent research in all its forms wherever it is found.
- To reduce significantly the administrative burden on institutions in relation to the RAE.
- To avoid creating any undesirable behavioral incentives.
- To promote equality and diversity.
- To provide a stable framework for our continuing support of a world-leading research base within HE.

The REF is a process of expert review. The assessed higher education institutes are asked to submit the required documents, publications, evidence concerning the outputs, impacts of the research

and research environment of the institute. The criteria for assessing outputs are rigor, originality, and significance. The criteria for assessing impacts are reach and significance or transformation. An expert sub-panel will assess the HEI's submission for each unit of assessment, working under the guidance of four main panels. Sub-panels will apply a set of generic *"assessment criteria and level definitions"* to produce an overall quality profile for each submission. The levels of the assessment criteria are five for each: Four star, Three Star, Two star, One star, and Unclassified. The expert panels of assessment will produce four quality profiles, with different weighting for each sub-profile awarded to each submission: Overall quality profile, Outputs sub-profile (65 percent), Impact sub-profile (25 percent), and Environment sub-profile (15 percent). The assessment criteria and level definitions of the four quality profiles are set out below.

The 2014 REF was conducted jointly by the Higher Education Funding Council for England (HEFCE), the Scottish Funding Council (SFC), the Higher Education Funding Council for Wales (HEFCW) and the Department for Employment and Learning, Northern Ireland (DEL). The REF was managed by the REF team, based at HEFCE, on behalf of the four UK higher education funding bodies, and was overseen by the REF Steering Group, consisting of representatives of the four funding bodies (REF 2014). The primary purpose of REF 2014 was to assess the quality of research and produce outcomes for each submission made by institutions (REF 2014):

1. The four higher education funding bodies will use the assessment outcomes to inform the selective allocation of their grants for research to the institutions which they fund, in effect from 2015-16.
2. The assessment provides accountability for public investment in research and produces evidence of the benefits of this investment.
3. The assessment outcomes provide benchmarking information and establish reputational yardsticks, for use within the higher education (HE) sector and for public information.

REF 2014 had taken years of toil and turmoil, but a new landscape of research excellence later emerged. An overall quality profile will be awarded to each submission to show the proportion of the submission that meets each of the starred levels outlined in Table 1.

Table 1. Overall Quality Profile: Definitions of Starred Levels

Four star	Quality that is world-leading in terms of originality, significance and rigor.
Three star	Quality that is internationally excellent in terms of originality, significance and rigor but which falls short of the highest standards of excellence.
Two star	Quality that is recognized internationally in terms of originality, significance and rigor.
One star	Quality that is recognized nationally in terms of originality, significance and rigor.
Unclassified	Quality that falls below the standard of nationally recognized work, or work that does not meet the published definition of research for the purposes of this assessment.

Source: REF (2014, p. 3).

Table 2. Outputs Sub-profile (65 percent)

Four star	Quality that is world-leading in terms of originality, significance and rigor
Three star	Quality that is internationally excellent in terms of originality, significance and rigor but which falls short of the highest standards of excellence.
Two star	Quality that is recognized internationally in terms of originality, significance and rigor.
One star	Quality that is recognized nationally in terms of originality, significance and rigor.
Unclassified	Quality that falls below the standard of nationally recognized work, or work that does not meet the published definition of research for the purposes of this assessment.

Source: REF (2014, p. 6).

Table 3. Impact Sub-profile (20 percent)

Four star	Outstanding impacts in terms of their reach and significance.
Three star	Very considerable impacts in terms of their reach and significance.
Two star	Considerable impacts in terms of their reach and significance.
One star	Recognized but modest impacts in terms of their reach and significance.
Unclassified	The impact is of little or no reach and significance; or the impact was not eligible; or the impact was not underpinned by excellent research produced by the submitted unit.

Source: REF (2014, p. 6).

The criteria for assessing the quality of outputs are "originality, significance and rigor" is depicted in Table 2, while the criteria for assessing impacts "reach" and "significance" are portrayed in Table 3. The research environment will be assessed in terms of its "vitality

and sustainability." Panels will consider both the "vitality and sustainability" of the submitted unit, and its contribution to the "vitality and sustainability" of the wider research base (see Table 4).

Table 4. Environment Sub-profile (15 percent)

Four star	An environment that is conducive to producing research of world-leading quality, in terms of its vitality and sustainability.
Three star	An environment that is conducive to producing research of internationally excellent quality, in terms of its vitality and sustainability.
Two star	An environment that is conducive to producing research of internationally recognized quality, in terms of its vitality and sustainability.
One star	An environment that is conducive to producing research of nationally recognized quality, in terms of its vitality and sustainability.
Unclassified	An environment that is not conducive to producing research of nationally recognized quality.

Source: REF (2014, p. 6).

In light of the importance of these standards in evaluating educational research, the Teaching and Learning Research Programme, in association with the Economic and Social Research Council (2009), conducted a review of the criteria used in the assessment of educational research across the UK. The Review Report *Quality Criteria for the Assessment of Education Research in Different Contexts*, was published in 2009. In addition to an account of general principles of research assessment, the report details three main contexts of evaluating educational studies: (1) peer review for publication, (2) Assessment of proposals for funding, and (3) investment decisions. The general principles used in research assessment are set out as follows:

1. Theoretical and methodological robustness, significance and contribution, and communication and engagement with different publics are general concerns that cut across, in different shapes, most contexts of assessment.
2. Operational criteria used in the assessment of education research may vary with the context, time, object, and purpose of assessment.
3. Research can be formative if it involves exchange, mutual learning and collaboration.
4. The actual review and assessment processes are based on individual and group interpretations of explicit criteria mediated by implicit standards of quality and worth.

The criteria used in the context of peer review for publication are classified into the following three categories:

1. **Criteria for journal publication:** relevance, significant and original contribution, accessibility and clarity, high standards of scholarship in argumentation and interpretation; sound methodological basis, ethics, adequate presentation, style, and language.
2. **Criteria fed into the assessment of monograph proposals:** balance of contents and coverage, convincing rationale and originality, logical structure, thematic coherence, and effective organization, overall academic standard, quality of writing of sample chapters, evidence of appropriate readership and market need, advantages over competition, suitability of author, peer endorsements, realistic timetable and feasibility, financial soundness, concision and clarity of proposal.
3. **Criteria used in the assessment of learned society conference:** scope, relevance, robustness, clarity, likely significance of presentation, appropriateness of research method and/ or literature, reporting sufficiently advanced research, accessibility to wider audiences, informative style, including adequate use of references and keywords.

The criteria used in the assessment of proposals for funding are classified in the following three categories:

1. **Criteria used in the assessment of proposals for funding research:** relevance, originality, innovation and topicality, scholarly importance, specificity, adequate background, clear, concise and appropriate objectives, aims and rationale for the project, explicit and appropriate theoretical and conceptual framework, promised scientific quality of approach, potential for positive users, society and economic importance and impact, user engagement, dissemination, contribution to capacity building, international competitiveness, ethical conduct and awareness of wider ethical conduct and awareness of wider ethical implications, principled scientific practice, ability, feasibility/achievability evaluation, value for money, presentation of the application, compliance with eligibility criteria and terms and conditions of the funding scheme.
2. **Criteria used in the assessment of applications to early career funding schemes:** outstanding academic merit of proposed project, merit of applicant, viability/feasibility of the project, suitability of the host institution, quality of plans for dissemination and engagement with academic and non-academic communities, potential impact, value for money, conformity to eligibility criteria and award specifications, quality of writing of the proposal.

3. **Criteria used in the assessment of proposals for the funding of developmental and practice-based research:** academic robustness, timeliness and potential to achieve positive short- or medium-term impact, educational usefulness, resonance, engagement of partners from different sites of education policy and practice in the research process, relevance to the remit of the funder, sustainability of development, locality, distinctiveness, ability/expertise of applicant to undertake the proposed research, professional development benefits for the applicant, institutional support, feasibility, evaluation, value for money, presentation.

The criteria of assessment related to investment decisions are classified in two categories:

1. **Criteria used in project and report evaluation:** accountability, effectiveness, quality, innovation, contribution to knowledge, contribution to policy and practice, contribution to research training and professional development, contribution to institutional development, development of research products, quality of output and effectiveness of dissemination, opening avenues for further research, peer appreciation and user satisfaction.
2. **Criteria used in the procurement of research:** conformance to the issued specifications and the associated terms and conditions, completeness, demonstrated understanding of the research brief, quality and credibility, added value, quality assurance, relevant leading-edge research expertise in the required discipline and substantive topic, relevant development experience, relevant strategic experience, communication and reporting, liaison and flexibility, organizational model, feasibility, consideration of the burden, consideration and adherence to ethical and legal issues, quality of management arrangements, value for money and soundness, price of contract.

Further important contexts for research assessment not included in the criteria are set out as follows: the awarding of degrees (MSc, PhD, EdD), the screening of existent research for review purposes, the inclusion of different forms of output in indexes, progress assessments/research monitoring, the evaluation of programs, the setting up of new journals or book series, the setting up of consultancy agreements.

Though these contexts are not covered in the proposed criteria of educational research quality of the Teaching and Learning Research Programme, they can apply these criteria in the process of research assessment, and in relation to different contexts. In short, the proposed criteria of ILRP are very universally inclusive and applicable to the evaluation of great varieties of genres of research documents.

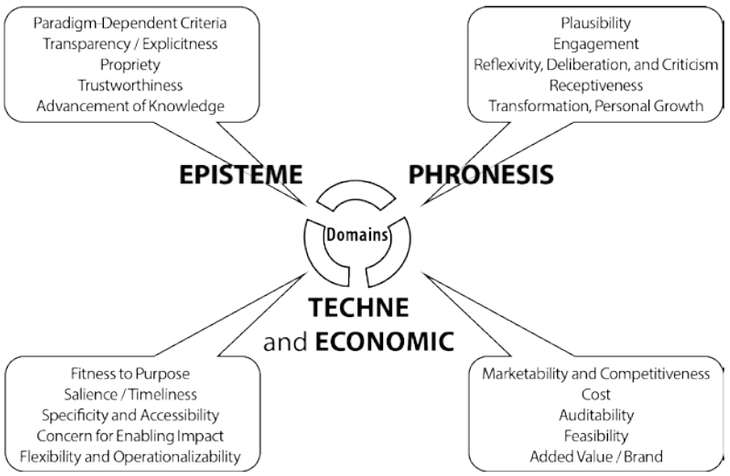
Inclusive and extensive as the proposed criteria of TLRP for the research assessment are, they mesh the theoretical with practical research in too simplistic a way and confuse practical with technological. To clarify the concept of quality and the relationship between theory, practice and technique, Furlong and Oancea (2005) conducted a research project *Assessing Quality in Applied and Practice-based Educational Research*, commissioned by the Economic and Social Research Council in 2004 and completed in 2005. In the research, they propose a three-dimensional approach for assessing educational studies. The three dimensions can each be divided into several sub-dimensions, which may be used to develop criteria for assessing the quality of research:

1. Epistemic: The assessment process should reflect traditional dimensions of quality such as methodological, theoretical and scientific robustness.
2. Technological: As immediate impacts of educational research on knowledge growth or improvement of policy and practice are not expected, the focus should be placed on the potential practical use, rather than just its actual impacts.
3. Capacity building and value for people: This dimension refers to the contribution to the collective and personal development of practitioners and policy-makers, making them more receptive to new ideas, encouraging new partnerships or improving their ability to reflect critically on their work.

The three-dimensional approach, each with some sub-dimensions, was elaborated and refined on the basis of Aristotelian distinctions between forms of rational activity and expressions of excellence on virtue. Three domains of excellence in applied and practice-based research are: theoretical (*episteme*), technical (*techne*), and practical (*phronesis*). The model of three domains of quality, each divided into sub-domains in applied and practice-based research, is epitomized in Figure 1.

The situation of educational research in Germany is much the same as in other countries. Confronted with the pressure of international competitiveness and being aware of the importance of educational research to national development, the German government and non-governmental agencies have recently initiated actions and measures to improve the infrastructure of educational studies and to enhance the quality of educational research.

Figure 1. Domains of Quality



Source: Adapted by the author from Furlong and Oancea (2007, p. 133).

After a series of workshops and negotiations, the Federal Minister of Education and Research and 16 State Ministers related to education reached a resolution, *Beschluss der Regierungschefs von Bund und Ländern zum Hochschulpakt*, based on which the federal government and the governments of 16 states (Länder) share the common responsibility for promoting the improvement of higher education institutes and for raising the standards for research in HEIs. This resolution was signed into a legal document in 2007 by the Federal and State Ministers with the title *Hochschulpakt 2020* ("*University Agreement 2020*"). In accordance with this agreement, the federal government lists the expenses for an extra-budget 70.3 billion Euro for improving research facilities of universities and supporting far-sighted research projects. Furthermore, based on the spirit of *Hochschulpakt*, the Federal and 16 State Ministers signed *Verwaltungsvereinbarung* ("Agreement of Administration"). Under the Agreement, the federal government lists the expenses for an extra-budget 14 billion Euro in 2010, 17.5 billion Euro in 2011, and 20 billion Euro each year from 2013 to 2020 to improve or recruit excellent and competent teaching and research personnel to forge ahead vigorously in improving the quality of teaching and research in HEIs.

In light of the importance of nongovernmental agencies in the enhancement of research quality, the federal government, in association with the 16 state governments in 2010 signed *Der Pakt für Forschung und Innovation* with Deutsche Forschungs-Gemeinschaft, Fraunhofer-Gesellschaft, Helmholtz-Gemeinschaft Deutshcer Forschungszentren, Max-Planck-Gessellschaft and Leibniz-Gemeinschaft. In accordance with the Agreement, the federal government promised an annual increment from 2011-2015 of 5 percent of the research budget for enhancing the international competitiveness of these research organizations. The concrete aims of this Agreement are as follows:

1. To promote the dynamic and progressive development of the German scientific research system.
2. To network all the German scientific research systems in order to enhance the capacity for research and development.
3. To develop or to improve strategies for international cooperation.
4. To build sustainable partnerships between scientific research and economic communities.
5. To continuously keep German scientific research quality at its apex. (BMBF 2010)

For the purposes of enhancing international cooperation and competitiveness in scientific research, on 15 May 2012, Wissenschaftsrat (2012) published *Leitfaden zur Begutachtung von Forschungsbauten-gültig ab Förderphase 2014*. In addidtion to the explanation of the aims, the composition of the assessment committee of the proposed guidelines for promoting the improvement of research construction, the Guideline promulgated a two-phase assessment (assessing an outline of the construction plan and a the details of the construction plan) and their evaluation criteria. The criteria for assessing the construction plan (both the outline and the details) are divided into five domains:

1. Objective-setting: Suitability of objective-setting for the building and required equipment for the applied research program.
2. Quality of the research program: Whether the superordinate scientific problem-raising of the applied research program is relevant, original and innovative, and whether the planned research work matches a coherent research program. How far is a committed medium- and long-term perspective expected in the proposed plan? How mature is the technical-scientific concept in the plan?
3. Quality of the preparatory work of the participating scientists in the plan: From their publications and past experience of participation or cooperation in a similar plan or project, is it possibble to show that their competency is suitable for carrying out the proposed plan?
4. The national significance of the proposed plan: How strong is the replication of the proposed plan? Can it be spread from one State (Land) to another? How important is it in national or international context? What position does the proposed plan have in relation to comparable research programs in Germany?

5. The inclusion of the proposed plan in the university: How does the proposed plan fit in the structural and development plans of the university, specifically, to the efforts of promoting up-and-coming academics, of parity-giving, of diversity management, of the transfer of science and technology, as well as the strengthening of competitiveness?

From the above discussions, it can be observed that the governmental research agencies in Germany are concerned not only with the research quality itself, but also with the construction and infrastructure for supporting high-quality research. Many independent non-governmental research agencies pay similar attention as governmental organizations, not only to research quality, but also to the required facilities for advancing high-quality research. The Leibniz Association (Leibniz-Gemeinschaft) is one such prominent research organization that not only produces excellent scientific studies, but also evaluates periodically the job, production, staff, management and quality-assurance of its affiliated institutions.

The Leibniz-Association, as one of larger non-university research organizations, can trace its origin to the Leopoldina established in 1652. In 1700, G.W. Leibniz initiated the founding of the Society of Sciences in Berlin, which was later to become the Academy of Sciences. After World War II, the federal political system was established, which empowers the federal states to control their own cultural affairs, education and science. However, some research institutions would have overburdened the funding capabilities of any one federal state. Thus, even before the Federal Republic was founded, the *Königstein Agreement* was signed at a meeting of West German federal states. In accordance with this agreement, joint funding for larger research institutions of super-regional importance was pledged if their financial needs exceeded the means of one individual federal state. The 1969 revised German Basic Law granted the Federal government and governments of States (Länder) the constitutional right to cooperate on research projects of supra-regional importance and national scientific interest. In 1977, after intensive negotiation involving more than 300 institutions, 46 were identified for joint funding and published, which was called the Blue List.

From 1979 on, the Science Council regularly evaluated Blue List institutions in order to guarantee high standards of scientific performance and be able to set them on the track to targeted development at an early stage. After German Reunification in 1990, the important research institutions were integrated into the Blue List. The total number of the institutions rose from 47 in 1989 to 81 in 1992. All 34 new members were accepted after successfully passing a Science Council evaluation. In 1990, the 81 institutions formed the Blue List Partnership, mainly for the purposes of cross-institutional administration. The List Partnership was transformed into the Blue List Science Association (WBL) in 1995. The WBL renamed itself the Gottfried Wilhelm Leibniz Science Association (WGL), known as the Leibniz Association for short. Currently, the Leibniz Association is composed of 89 Institutes belonging to five sections.

The Section of Humanities and Educational Research has 11 institutes (five for educational science and six for cultural and historical science) and five museums. The main tasks of this Section are: to carry out scientific research on cultural reality and construction and to promote critical analysis of scientifically acquired cultural knowledge; to document cultural heritage and to guarantee its restoration; and to disseminate the results of scientific research and documentation to both the professional and the non-professional publics.

In order to guarantee the fulfillment of the mentioned tasks and the quality of their performance, the Science Council launched a seven-year periodic evaluation of the institutes of the Leibniz Association beginning in 1979. The evaluation procedure is divided into two stages: the first stage is conducted by review boards, the second by the Leibniz Association Senate.

The criteria of the first stage evaluation are set out in the following categories:

1. General concept and profile/subdivisions of the institutions: a coherent and successive development plan, originality and state-of-the-art work program, relevance of the institute's work to the non-academic public, a strong position in national and international arenas, the number of publications, commercial property rights and patents, consulting contracts and third-party funds raised for research, visibility and outreach of the institute's performance, and appropriateness of staffing and facilities.
2. Collaboration and networking: commitment to collaboration with universities, joint academic appointments with universities, appropriate cooperation with non-university institutes at home and abroad.
3. Staff development and promotion of junior staff: appropriate strategies to recruit personnel, good structure of the composition of personnel, staff development and continuing education, gender equality, structured programs for promoting junior academic staff, professional training for non-academic staff.
4. Quality assurance: the appropriateness of the leadership and management staff, the appropriateness of in-house quality assurance measures, the reasonableness of the composition of the scientific advisory board/user advisory board and the supervisory board, and the effectiveness of the board's work.

The criteria for science policy-related evaluation by the Leibniz Association Senate (the second stage of the evaluation procedures) include quality of output, excellence of work units, achievement potential of the institution, added value of the institution, the significance of the institute's strategies and the international visibility of the institute's achievements.

Not only government agencies and independent corporations, like the Leibniz Association, but also professional academic learned societies, have launched initiatives to improve educational research quality and development. Fretting over the lack of undisputed quality-standards and of international connection, *the Zeitschrift f ür Erziehungswissenschaft* held a two-day ZfE-Forum on 12-13 December 2003 in Berlin with the theme "Standards und Standardisierung in der Erziehungswissenschaft." The Forum was intended for discussions on the quality-deficiency in various strategic points of education and educational science, on the formulation of standards, on the ways and process to actualize the formal standards of education and educational science (Gogolin et al. 2005). The results of the discussions in the Forum were published as *Beiheft* der Zeitschrift *f ür Erziehungswissenschaft* in 2005.

It is worth noting that a document for evaluating research and teaching at universities proposed for the discussions in the Forum by the Ministerium für Wissenschaft, Forschung Und Kunst Baden-Württemberg (2004) manifests as typical concerns about the standards of university research and teaching.

The criteria for assessing research quality proposed by the Baden-W ürttemberg Ministry are divided into the following eight categories:

1. Fundamental conditions include: material construction (research space, technical infrastructure, laboratory, library facilities, financial resources) and personnel provisions (the size and range of qualifications profiles of personnel with academic expertise, the relationship of research personnel to faculty members of the university, the number of student assistants, technical personnel, and administrators and office personnel).
2. Scientific quality of the research is evaluated in terms of the following criteria: a clearly designed research program, scientific profile-formulation and discernibility of the research focus, originality of approaches and problem-formulations, coherence and cumulativeness of the research results, the significance of the research in national and international development of the studied field, recognition by the scientific community (invitations to national and international conferences, editors of renowned scientific journals, reviewers of the German Research Association and international research organizations, and prominent position in the scientific system).
3. The criteria for assessing production quality of research in HEIs are as follows: academic publication (the number of articles published in renowned, double blind-reviewed journals, books or book chapters, monographs in renowned publication series or university publication series), number of dissertations, contributions to international or national conferences, professional publications for practice, policy-making or the general public, and patents or productions for further development.
4. Practical relevance of research is assessed in terms of: its relevance to the dissemination of professional knowledge, relevance to educational practice through the implementation of innovations and the improvement of and relevance to educational decision-making and public discussions.
5. The enlistment of third-party funds is assessed in terms of winning competitive funds offered by research-promotion organizations or enlisting other financial resources, the systematic promotion of the younger-generation research staff with third-party funds, and earning production through third-party funds.
6. The support given to younger-generation research is evaluated in terms of the following aspects: the embodiment of the younger works for applying for qualification in the research projects and systematic care for younger researchers, a plan for the professional development of younger-generation researchers through methodical education and participation in conferences, joint publications of junior and senior scientists, participation of external evaluators, and active participation in the graduate program.
7. Cross-regional cooperation and international connection are evaluated in terms of the following aspects: participation in coordinated cross-regional programs of the German Research Association, successfully winning the resources provided by the EU, Federal or State Governments or Foundations that initiate research programs, a cross-regional or regional developmental plan as the focus of the research, exchanges of guest researchers financed by DAAD or other research-supporting agencies, and participation in international research societies.
8. The management and developmental plan is assessed in terms of whether the HEIs have a systematic plan for quality assurance and quality development for sustainable growth of the institutes.

From the foregoing comparison of educational quality-standards in international organizations and various countries, the following

common characteristics of high-quality standards construction can be observed:

1. Attention is paid, not only to the construction of the standards of good quality itself, but also to how the hardware and infrastructure favorably contribute to conducting excellent educational research.
2. Rigor, originality, significance, integrity and style of communication are the most important common indicators employed in the assessment of educational research quality in various countries.
3. Ethical considerations are specifically emphasized both in the process of conducting research and in the dissemination and utilization of the studied results.
4. The impacts of educational research are evaluated, not only in light of immediate effectiveness, but also considering its possible long-term influences.
5. The outreach of the impacts should be evaluated in terms of the influences upon educational decision-making by practitioners, as well as by all the stakeholders of educational research.
6. The assessment of the quality of research projects should take into consideration their embedment into the structure and programs of HEIs, especially into plans for enhancing research capacity and promoting opportunities for younger-generation researchers.
7. Innovation and advancement of educational knowledge are specifically emphasized in the assessment of the highly competitive apex research projects provided by international organizations or various governmental or nongovernmental agencies.
8. The assessment of educational research should vary from context to context, depending on its different users (e.g., journal publication, selection of monograph series or promotion of professional positions).
9. International visibility of educational research and development is emphasized in various countries.

Concluding Remarks

Educational research has been criticized as sloppy, useless, fragmented, irrelevant to decision-making and practice, and lacking cumulativeness for the growth of knowledge, and, even worse, as wasting limited national resources doing insignificant, idle things for a long time in international organizations and various countries.

Accompanying the quick progress of science and technology and the rapid pace of globalization, international competitiveness has led to mounting tensions day by day. Being aware of the important role of high-quality education supported by excellent research for national development, international organizations and many countries in the world have established institutes with special responsibility for monitoring the quality and progress of educational research.

The research institutes established by international organizations and governments in various countries are generally commissioned not only with the task of monitoring research quality, but also with the mission to improve research facilities and conditions for recruiting and promoting younger-generation researchers. The Centre for Educational Research and Innovation of OECD, the Joint Research Centre of EU, the Office of Educational Research and Improvement in the United States, the Max-Planck-Institut für Bildungforschung and the National Foundation for Educational Research in the UK are all responsible, not only for monitoring research quality itself, but also for improving research environments to facilitate the production of high-quality educational research.

One of the important measures employed by these institutes to support high-quality educational research is research assessment. For the purposes of conducting fair and objective assessment, almost all of these institutes have developed their own quality standards. The most important common standards for assessing research quality itself are rigor, originality, significance, integrity and communication style. However, the detailed and concrete criteria of each standard are under debate by different sectors involved in educational research.

It is therefore imperative to strengthen international cooperation of international, national and independent educational research organizations. Through the coordination and cooperation of international, national and professional educational research agencies, a fair and objective series of criteria for judging educational research quality can be developed. Under the assessment of the newly-developed criteria, the improvement of educational research can be expected, thus leading to the betterment of educational practice. Since education plays a key role in social, cultural and economic development, high-quality education activities supported by excellent educational research may lead to elevating the happiness of all humanity living in an environment of flourishing development.

References

American Educational Research Association (AERA). 2006. "Standards for Reporting on Empirical Social Science Research in AERA Publications." *Educational Researcher* 35 (6): 33-40.

AERA. 2009. “Standards for Reporting on Humanities-oriented Research in AERA Publications.” *Educational Researcher* 38 (6): 481-486.

AERA. “Code of Ethics.” Educational Researcher 40 (3): 145-156.

Andresa, Gruschka. 2013. “‘Empirische Bildungsforschung’ am Ausgang ihrer Epoche?” *Profil* (June): 1-8.

Bilstein, Johannes. 2009. “Standardisierung-Kanonisierung. Zur Imaginations und Begriffsgeschichte.” In *Standardisierung-Kanonisierung. Erziehungs-wissenschaftliche Reflexionen*, edited by Johannes Bilstein and Jutta Ecarius (pp. 15-29). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Blossfeld, Hans-Peter, Jörg Doll, and Thorsten Schneider. 2009. “Methodological Advantages of Panel Studies Designing the New National Educational Panel Study (NEPS) in Germany.” *Journal for Educational Research Online* 1 (1): 10-32.

Blossfeld, Hans-Peter, Jutta von Maurice, and Thorsten Schneider. 2011. *Grudidee, Konzeption und Design des Nationalen Bildungspanels für Deutschland*.” NPES Working Paper 1. Bamberg, Germany: Federal Ministry of Educationa and Research.

Bottle, Alexander. 2007. “Scientometric Approaches to Better Visibility of European Educational Research Publications: A State-of-the-Art Report.” *European Educational Research Journal* 6 (3): 303-311.

Bridges, David. 2009. “Assessing the Quality of Research in Higher Education: The UK Research Assessment Exercise.” In *Assessing the Quality of Educational Research in Higher Education International Perspectives*, edited by Tina Besley (pp. 85-104). Dordrecht, Netherlands: Sense Publishers.

British Educational Research Association. 2013. *Why Educational Research Matters: A Briefing to Inform Future Funding Decisions*. London: British Educational Research Association.

Brody, Tim. 2014. *Citation Analysis in the Open Access World*. Retrieved from http://eprints.soton.ac.uk/id/eprint/260000; accessed on 13 June 2014.

Campbell, David F. J. 2005. *Evaluation of University Research in Europe: Conceptual Framework and Comparative Typology*. Vienna: Faculty for Interdisciplinary Studies.

Coe, Robert. 2013. *Improving Education: A Triumph of Hope Over Experience*. Durham, UK: Centre for Evaluation and Monitoring, Durham University.

European Commission. 2010. *Simplifying the Implementation of the Research Framework Programmes*. Brussels: European Commission.

European Commission. 2012. *Global Europe 2050*. Luxembourg: Publications Office of the European Union.

European Commission. 2015. *HORIZON 2020*. Brussels: European Research Council.

Feyfant Annie, and Rey Olivier. 2006. *What is “Good” Research in Education?* La lettre d'information No. 22. Lyon, France: École normale supérieure de Lyon.

Fletcher, M. Jack, Lisa Towne, and Lauress L. Wise. (eds.). 2004. *Strengthening Peer Review in Federal Agencies that Support Education Research*. Washington, DC: The National Academies Press.

Gogolin, Ingrid, Heinz-Hermann Krüger, Dieter Lenzen, and Thomas Rauschenbach. 2005. “Editorial. In *Standards und Standardisierung in der Erziehungswissenschaft*, edited by Gogolin, Krüger, Lenzen and Rauschenbach (pp. 5-10). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Gough, Stephen, and William Scott. 2007. “Education and Sustainable Development: A Political Analysis.” *Educational Review* 58 (3): 273-290.

Gregg. B. Jackson, and Richard C. Atkinson, eds. 1992. Research and Education Reform: Roles for the Office of Educational Research and Improvement. Washington, DC: The National Academies.

Gregor, Lang-Wojtasik. 2005. “Qualität, Qualitätsmessung und Qualitätssicherung im Kontext von Education For All.” *Zeitschrift Für internationale Bildungsforschung und Entwicklungspädagogik*,28: 29-30.

Gretler, Armin. 2007. “The International Social Organisation of Educational Research in Europe: Reviewing the European Educational Research Association as an Example – Facts and Questions.” *European Educational Research Journal* 6 (2): 174-189.

Harrison, Jayne E., and Sandeep Sharma. 2006. “Structured Abstracts: Do They Improve the Quality of Information in Abstracts?” *American Journal of Orthodontics & Dentofacial Orthopedics*130 (4): 523-530.

Harvey, Lee, and Diana Green. 1993. “Defining Quality.” *Assessment and Evaluation in Higher Education* 18 (1): 9-34.

Harvey, Lee. 1995. “Editorial: The Quality Agenda.” *Quality in Higher Education* 1 (1): 5-12.

Higher Education Funding Council for England (HEFCE). 2009. *Research Excellence Framework: Second Consultation on the Assessment and Funding of Research*. Bristol, UK: HEFCE.

HEFCE. 2014a. *Assessment Criteria and Level Definitions*. Bristol, UK: HEFCE.

HEFCE. 2014b. *Independent Review of the Role of Metrics in Research Assessment: Call for Evidence*. Bristol, UK: HEFCE.

HEFCE. 2014c. *Policy for Open Access in the Post-2014 Research Excellence Framework*. Bristol, UK: HEFCE.

Hillage, Jim, Richard Pearson, Alan Anderson, and Penny Tamkin. 1998. *Excellence in Research on Schools*. DfEE Research Report No 74. London: Department for Education and Employment.

Hood, Paul. 2002. *Perspectives on Knowledge Utilization in Education*. San Francisco, CA: WestEd.

Institut für Hochschulforshung Wittenberg. 2009. *Politische Steuerung und Hochschulentwicklung unter föderalen Bedingungen*. Wittenberg, Germany: HoF-Arbeitsbericht.

Institute of Education Sciences (IES). 2016b. *Procedures for Peer Review of Grant Applications*. Washington, DC: IES.

IES. 2006a. *Procedures for Peer Review of Reports*. Washington, DC: IES.

IES. 2008. *Rigor and Relevance Redux: Director's Biennial Report to Congress* (IES2009-6010). Washington, DC: U.S. Department of Education.

Kaestle, Carl F. 1993. "The Awful Reputation of Education Research." *Educational Researcher* 22 (1): 23, 26-31.

Kerlinger, Fred N. 1960. "The Mythology of Educational Research: The Methods Approach." *School and Society* 88: 149-151.

Kerlinger, Fred N. 1977. "The Influence of Research on Educational Practices." *Educational Researcher* 6 (8): 5-12.

Lagemann, Condliffe Ellen. 2000. *An Elusive Science: The Troubling History of Education Research*. Chicago: University of Chicago Press.

Leibniz Association. 2012. *The Senate. 2012. The Leibniz Association Senate Evaluation Procedure: Basic Principles*. Berlin: Leibniz Association.

Manning, Sabine. 2007. "Linking Educational Research Activities across Europe: A Review of the WIFO Gateway to Research on Education in Europe." *European Educational Research Journal* 6 (4): 446-450.

Martin, Ben R. 2011. "The Research Excellence Framework and the 'Impact Agenda': Are We Creating a Frankenstein Monster?" *Research Evaluation* 20 (3): 247-254.

Martina, R. (bearb). 2004. *Evaluation der Erziehungswissenschaft an den Universitäten und Pädagogischen Huchschulen des Landes Baden-Württemberg*. Final Report. Bonn: Ministerium für Wissenschaft und Kunst.

Oancea, Alisa. 2011. "How Far Have We Come? The Current State of Education Research." *Research Intelligence* 115 (6).

Oancea, Alisa, and John Furlong. 2007. "Expressions of Excellence and the Assessment of Applied and Practice-based Research." *Research Papers in Education* 22 (2): 119-137.

Organisation for Economic Co-operation and Development (OECD). 1995. *Educational Research and Development: Trends, Issues, and Challenges*. Paris: OECD.

Ozga, Jenny. 2008. "Governing Knowledge: Research Steering and Research Quality." *European Educational Research Journal* 7 (3): 261-272.

Pollard, Andrew. 2008. *Quality and Capacity in UK Education Research*. Forum I Report. Harrogate, UK: Strategic Forum for Research in Education.

Prost, Antoine. 2001. *Pour un Programme Strategiqué de Recherche en L'éducation: Raport du Groupe de Travail Constitué par Antoine Prost*. Paris: Ministére de l'Éducation Nationale.

RAND Foundation. 2011. *RAND's Standards for High-Quality Research and Analysis*. Los Angeles, CA: RAND Corporation.

RAND Corporation. 2013. *Standards for High-Quality Research and Analysis*. Available online at: http://www.rand.org/standards.html; accessed 2 March 2017.

Ranis, Sheri H., and Walters Pamela Barnhouse. 2004. "Education Research as a Contested Enterprise: The Deliberations of the SSRC-NAE Joint Committee on Education Research." *European Educational Research Journal* 3 (4): 795-806.

Reeves, Thomas. 2011. "Can Educational Research be Both Rigorous and Relevant?" *Journal of the International Society for Design and Development in Education* 1 (4): 1-24.

Research Excellence Framework (REF). 2011. *REF 2014: The Results*. Available online at: http://www.ref.ac.uk/pubs/201401; accessed 2 March 2017.

Koul, Rekha B. 2008. "Educational Research and Ensuring Quality Standards." *E-journal of All India Association for Educational Research (EJAIAER)* 20 (3&4).

Rey, Olivier. 2011. "An Overview of the Current Position of Education Research in France." *Research Intelligence* 115: 28-29.

Richard, Van Heertum. 2005. "How Objective is Objectivity? A Critique of Current Trends in Educational Research." *UCLA Journal of Education and Information Studies* 1 (2): 1-13.

Royce, Josiah. 1891. "Is There a Science of Education?" *Educational Reviews* I: 23-24.

Schuller, Tom. 2005. "Constructing International Policy Research: The Role of CERI/OECD." *European Educational Research Journal* 4 (3): 170-180.

Shavelson, Richard J., and Lisa Towne (eds.). 2002. *Scientific Research in Education*. Washington, DC: National Academy Press.

Tharp, Roland G. 1997. *From At-risk to Excellence: Research, Theory, and Principles for Practice*. Research Report No. 1 Berkeley, CA: Center for Research on Education, Diversity & Excellence.

Tooley, James. 2001. "The Quality of Educational Research: A Perspective from Great Britain." *Peabody Journal of Education* 76 (3/4): 122-140.

von Maurice, Jutta, Thomas Leopold, and Hans-Peter Blossfeld. 2011. *The National Educational Panel Study: A Long-Term Assessment of Competence Development and Educational Careers*. NPES Working Paper 2. Bamberg, Germany: Federal Ministry of Educationa and Research.

Wissenschaftsrat. 2008. Empfehlungen zur Einrichtung einer programmatisch-strukturellen Linie "Hochleistungsrechner." *In Rahmen der Förderung von Forschungsbauten an Hochschulen einschließlich Grogeräten,* nach Art. 91b Abs. 1 Nr. 3GG. Berlin: Wissenschaftsrat.

Yates, Lyn. 2005. “Is Impact a Measure of Quality? Some Reflections on the Research Quality and Impact Assessment Agendas.” *European Educational Research Journal* 4 (4): 391-403.