Measuring by Numbers: Bibliometric Evaluation of Faculty's Research Outputs and Impact on Academic Life in China

Wenqin Shen, Dan Mao, and Yaqiong Lin

Introduction

The use of quantification as a management tool is evident across the world (Shore and Wright 2015; Muller 2018). Government agencies are adept at using statistics and numbers to aid administration. In the field of higher education, the use of quantitative management on academics and their research activities worldwide has been increasing (Brenneis et al. 2005). This trend is motivated by various reasons, which include the rampant development of science and technology and the democratization of higher education (Weingart 2010).

In the UK, an increasing number of university teachers find themselves the subject of performance appraisals, which are inherently in conflict

W. Shen (✉)
Peking University, Beijing, China
e-mail: shenwenqin@pku.edu.cn

D. Mao
Shanghai Normal University, Shanghai, China

Y. Lin
Huazhong University of Science and Technology, Wuhan, China

203

with the institutional logic of the university (Townley 1997). In countries such as Australia, research performance evaluation has become a regular regime for faculty. One Australian university evaluates its faculty members on their performance in obtaining external funds, supervising Ph.D. students, and publishing (Whitley et al. 2010; Welch 2016). According to interviews conducted, teachers at the History Department of one Australian university needed to report the number of received citations by their publications when applying for promotion (Gläser and Laudel 2007). In France, the evaluation of teachers varies among different types of universities. However, the practice of assessing faculty research performance by journal classification and number of citations still exists (Paradeise and Thoenig 2015). Norwegian scholars who have analyzed research output by teachers at different age cohorts discover output for all cohorts in an upward trend, which is linked to university incentives (Kyvik and Aksnes 2015).

Traditionally, American research universities tend to rely less on quantitative approaches in their teachers' research appraisals. Such case is particular with top research universities; for example, neither UC Berkley nor MIT sets clear quantitative standards or requirements on the number of publications when promoting or recruiting teachers, and scholars' citation statistics in evaluation reports are for reference only (Thoenig and Paradeise 2014). In an interview by an American researcher, the head of the Chemistry Department of an American university pointed out that their rules on faculty promotion are "vague;" "In our faculty promotion guide, you won't find requirements that you have to have three research funds or publish six papers" (Nadler 1999, p. 61). However, other evidence suggests that even in the United States, the number of publications is becoming increasingly important in academic promotion. Many young teachers are informed that the primary factor that decides promotion is the number of publications (Anderson et al. 2010). Contrary to traditional scientific and sociological theories, analysis of the academic output by American scholars in social sciences reveal that an increasing number of scholars are publishing peer-reviewed papers, followed by a sharp increase in the scale of academic output. Such an increase can be caused by the encouragement to publish on the part of universities and the link between additional publications and high incomes (Hermanowicz 2016).

Although research publications by Chinese scientists have surged over the past two decades (Liu et al. 2015), it is less impressive in qualitative terms than quantitative. Studies show that even compared with top

universities in Hong Kong and Taiwan, research universities in Mainland China generate research output in larger quantity but of lower quality (Li et al. 2011). The present work argues that one reason behind the surge in publications by Chinese scholars is to cope with the quantitative appraisals by their institutions. The impact of bibliometric evaluation or audit culture on Chinese university teachers has caught the attention of some researchers (e.g., Yi 2011). Other scholars have analyzed the implications of global university rankings or national research evaluation regimes on university faculties (Li 2016), and others still studied the reaction of Chinese university teachers to new managerialist reforms (Huang et al. 2018).

Generally, existing research has yet to offer in-depth analysis dedicated to bibliometric evaluation and its implications, leaving a few questions unanswered. First, current research focuses on the policy implications at a global or national level (Li 2016), providing little attention to the variety or discrepancies among Chinese universities. Second, existing research examines the influence of research evaluation systems on Chinese university teachers but has not provided due attention to the impact of a special evaluation approach (i.e., bibliometric evaluation). Last, further attention has been provided to the natural sciences than social sciences, wherein making an equivalent impact has been proven complex for Chinese scholars. Therefore, this study seeks to analyze the quantitative management in the research evaluation of Chinese teachers and its implications for the research life of Chinese academics by interviewing 36 university teachers from 8 universities and analyzing relevant policy texts and bibliometric data from a selected Project 985 university.

The Authors interviewed scholars of engineering, physics, chemistry, Chinese literature, history, and sociology to present the impact of bibliometric research evaluation on faculty members of different disciplines in a balanced fashion. The interviews were conducted over an extended period in 2004, 2007, 2011, and 2015, which allowed us to observe historical continuity and changes. Output data for researchers in the two fields of education and anthropology in one Project 985 university were also analyzed for the years 1993, 2003, and 2013 to measure change over time.

Moreover, we assembled relevant policy texts on recruitment, research reward, professional promotion, workload appraisal of 10 research universities, and the "development schemes for the 12th and 13th Five-Year Plans" of 75 universities directly under the Ministry of Education to

understand how quantitative research evaluation is reflected in the policy texts of various universities.

This chapter is structured as follows. First, we present how quantitative evaluation, as a social technology, historically entered and gained legitimacy in the domain of higher education by reviewing relevant texts. Second, we illustrate how quantitative metrics evaluation is reflected in university policies by analyzing policy texts of individual universities. Finally, we analyze how national and university quantitative evaluation policies influence university teachers' academic lives, decisions regarding how and where to publish, and particular knowledge production activities.

The Rise of Bibliometric Evaluation in Chinese Universities

In the 1980s, quantitative evaluation had not gained prevalence in faculty evaluation in Chinese universities. In appraisals for promotion to associate professorship and professorship, seniority was part of the consideration, as well as reputation and influence in the field. In other words, peer review played a rather important role at that time. For example, in the 1980s, professorship appraisals in Peking University entailed reporting and defense at the university's academic council (Interviews with members of Peking University's academic council 2012).

From the end of the 1980s onward, bibliometric evaluation methods were gradually introduced first into natural sciences. In 1987, as required by the New Technology Bureau of State Scientific and Technological Commission, the Institute of Scientific and Technical Information of China (ISTIC) conducted statistical analysis of Chinese scientific publications between 1983 and 1986 indexed by SCI, ISR, and ISTP using bibliometric methods. In 1988, Shang Yichu of ISTIC published the top 10 Chinese universities in terms of academic publications from 1983 to 1986 in the report *China's Academic Standing in the World*. Since 1988, the Department of Science and Technology of the Ministry of Education has regularly published the Compilation of Science and Technology Statistics of Higher Education Institutions, which compiled data on researchers, research funds, and publications of various universities. Such official data served as reference for the comparison between the numbers of scientific papers published by different universities. Data for 1988 show that the number of academic publications by Chinese universities was low, with Peking University having 2412 researchers, publishing 1299 papers

or 0.54 paper per researcher. Equivalent numbers for Tsinghua University were 5768, 1293, and 0.22 (National Education Commission of the People's Republic of China 1989).

In 1989, commissioned by the Department of Comprehensive Planning of the State Scientific and Technological Commission, ISTIC conducted statistical analysis of scientific papers published in Chinese language journals. Since then, annual statistical analysis on the papers indexed by the three major international indexes for scientific literature (SCI, ISR, and ISTP) has become a regular endeavor. In 1990, the Chinese government evaluated National Key Labs. Subsequently, the Chinese Academy of Sciences appraised its subordinate research institutes, providing considerable attention to the number of publications. In 1992, for the first time, Nanjing University overtook Peking University and became number one in the list of SCI-indexed papers. In 1993, Nanjing University still held the first place, followed by Peking University (China Institute of Scientific and Technical Information 1995). On October 23, 1998, the Ministry of Science and Technology held a press conference, where it officially published the Statistics on Chinese Scientific Publications in 1997, which ranked Chinese universities. This report received a considerable amount of attention from university leaders across the country, being from such an official source.

Since then, SCI rankings have become increasingly important among Chinese universities. Moreover, with a policy orientation heavy on SCI papers, some traditional engineering institutions began to emphasize the number of SCI papers.

In 1998, the central government launched the Project 985, which aimed to build some world-class and top research universities. In 2002, the Ministry of Education officially launched First-level Discipline Rankings, with many of the evaluation metrics being bibliometric. Zhejiang University issued Provisions on Thesis Defense for Postgraduate Degrees of Zhejiang University in the same year, requiring a certain number of publications before Ph.D. students' thesis defense. Peking University would later adopt this practice, with other universities soon following suit.

In 2003, Shanghai Jiao Tong University presented the world's first Academic Ranking of World Universities, which is a thorough presentation of research performance that has since become highly important for universities worldwide. Some unofficial university rankings, such as Wu Shulian's Chinese University Ranking, have also been adding pressure on

universities. Therefore, 2002–2003 is a key milestone in the history of higher education in China, especially regarding research evaluation.

Driven by many university rankings, universities generally provided considerable importance to the number of scientific papers. In a 2004 interview, one teacher identified university rankings as one driver behind universities' emphasis on publishing additional papers:

> This is not easy. This decides a university's ranking. So the university leaders are very nervous about this, so they put high requirements on students. I don't know if they have this kind of requirements outside China. But for our university, they are always stressed about rankings. And there's quite a gap between us and universities in Beijing and Shanghai. So the university requires publications. And it really worked and our ranking went up. (Interview, Professor at a Project 985 university 2004)

With regard to the introduction of the World-Class University Project, the number of SCI papers has become a key metric pursued by university leadership. When some universities summarize their achievements, their ranking in the number of SCI and EI indexed papers in China is commonly featured. For example, when a Project 985 university summarized its successes, it indicated that their SCI and EI paper ranking among Chinese universities rose from 19th and 14th in 2009 to 7th and 8th in 2013, respectively. On this basis, numbers of papers, impact factor of the journal, and number of citations have been widely used in the assessment of teachers. Bibliometrics is an important criterion in annual performance assessments and academic promotions. Furthermore, bibliometrics has become a key criterion for important academic rewards. Many academic rewards, such as Ministry of Education's Award in Research Achievements and Cheung Kong Scholar Program, require candidates to submit citation statistics of their papers.

We selected a Sociology Department and a College of Education faculty members in one Project 985 university as cases to illustrate the historical construction of quantitative assessment in Chinese universities. We analyzed the research publication data of the two faculty members in 1993, 2003, and 2013. We found that the scale of teachers in both schools has considerably increased after the implementation of the Project 985 and the massification of higher education. The number of faculty in the Department of Sociology has increased from 4 in 1993 to 11 in 2003 and further expanded to 28 in 2013. The College of Education has begun

Table 1 Total number of journal articles/number of teachers

	1993	*2003*	*2013*
Department of Sociology	6/4	33/11	92/28
College of Education	10/12	79/21	61/27
Total	16/16	112/32	153/55

Table 2 Number of published articles in Chinese journals per capita

	1993	*2003*	*2013*
Department of Sociology	1.5	3.0	3.3
College of Education	0.8	3.8	2.3
Total	1.0	3.5	2.8

to take shape in 1993 and has 12 teachers. It expanded to 21 in 2003 and further increased to 28 in 2013 (Table 1).

In comparison with 1993, the per capita publication number of teachers in the two units showed a rapid development trend in 2003. The per capita publication number of teachers in Department of Sociology indicates a continuous growth trend. From 1993 to 2003, the per capita number of journal articles published by the members of the Department of Sociology doubled from 1.5 to 3.0; from 2003 to 2013, this figure continued to increase, although the growth rate slowed down to only 10% (Table 2).

In addition, the two colleges do not particularly emphasize the English publication in their faculty evaluation system; thus, the number of English publications has been limited in the two colleges. In the College of Education, the number of English publications was 0 in 1993, 3 in 2003, and 4 in 2013. In the Department of Sociology, no teachers published English papers in 1993, 2003, and 2013.

Bibliometric Evaluation in University Policies

Across the board, Chinese universities increasingly embrace bibliometric evaluation as a tool to manage teachers' research performance. As a management method, quantitative appraisal has permeated into every facet of faculty performance assessment, including university development planning, performance assessment, research incentive and academic promotions, thereby constructing an institutionalized system based on official policy texts.

University Strategic Planning

Soon after the founding of the People's Republic, planning systems were introduced into Chinese universities. Development planning is a tool for macro control by the state and an important way of autonomous governance by universities (Qi and Chen 2016). It is a compass that guides an institution's development in the coming years and exerts major influence on its various management systems and policies. Via an documentary analysis of the 12th and 13th Five-Year Plans of 75 universities directly under the Ministry of Education, it is found that setting quantitative metrics is common in the research development plans of universities. For example, a Project 211 university in Jiangsu Province indicated specific expectations as the following in its 13th Five-Year Plan:

> By the end of the 13th five year, we will strive to have 500 thousand RMB in per capita research funding for faculty members with senior titles, with total of 3 billion RMB research funding in place for the university. Among which, total horizontal funding (from private companies) shall exceed 500 million RMB… strive to secure at least 8 new important national scientific projects or key projects from key development programs, an average of 110 million RMB of project funding for National Natural Science Foundation of China projects, 15 key important projects from National Social Science Fund of China, 45 projects from National Social Science Fund of China, 5–6 international collaboration projects… 6 new National Science and Technology Awards, strive to achieve breakthrough in State Science and Technology Prizes, 2 new Award in Research Achievements in Humanities and Social Sciences of the Ministry of Education. Publish at least another 40 high-level (IF $\geq$ 9)) SCI papers, 7000 SCI/EI papers, 100 SSCI papers and 1800 CSSCI papers in total.

As a key metric for various evaluations and rankings, research achievement has become a crucial means to boost university ranking and reputation with the continued implementation of the World-Class University vision. The comparison of the research results during the 12th and 13th Five-Year Plans revealed that some universities not only named specific quantitative targets on research funding and scientific publications while formulating their 13th Five-Year Plan but also proclaimed high targets for the growth rate of these metrics. Consequently, a finance and economics university which had published 75 SCI papers and 90 SSCI yearly on average during the 12th five-year period boosted the numbers

of SCI and SSCI publications by over 100% to above 200% during the 13th five-year period.

Some universities began using key performance indicators in their planning and management to achieve the research performance targets set out in the 13th Five-Year Plan. Furthermore, they devolved the metrics level by level, all the way down to individual faculties and departments. For example, a Project 211 university in Central China listed the main performance metrics in research in its Notice on the Proposed Targets for 2017: National Key Projects (e.g., important and key projects from the National Natural Science Foundation), major government awards (e.g., State Science and Technology Prizes), and high-level papers (e.g., SCI and SSCI papers). When setting annual targets, schools and departments shall make "targets that are quantifiable, assessable with visible results, and weighted" (CCNU 2016). In many universities, the quantity of research publications is the core metric used in the annual assessment of individual schools and departments. The development planning texts of various universities indicate that the number of academic papers is still the primary target that many universities strive to meet or exceed. The quality of a paper is mainly determined by whether the journal in which it is published has been included in major bibliometric indexes (e.g., SCI, EI, SSCI, and CSSCI); thus, judging the quality of the paper is simply by the journal's impact factor.

Carrots for Publication

As mentioned earlier, numerous universities have planned to boost the number of research publications under the pressure of building world-class universities. However, if these plans are to be fully implemented at the level of individual teachers, a series of support policies are needed. Therefore, teachers and departments tend to heavily rely on rewards for scientific publications and research projects.

Generally, rewards for research performance mainly include the following types: (1) national-, provincial-, or ministerial-level research awards; (2) research papers or monographs; and (3) national-, provincial-, or ministerial-level research projects. Table 3 shows the research reward standard of a Project 985 university in Western China, such as rewarding of each paper with 30,000 RMB based on the journal of publication; each award for outstanding research achievement at national-, provincial-, or ministerial-levels ranges from 50,000 to 2 million RMB.

Table 3 Research reward standard of a project 985 university in Western China

Order	Category		Reward standard (thousand RMB)
Reward standard for academic papers in natural sciences			
1	Published in Science or Nature		300
2	Published in sister journals (research journals) of Science or Nature		100
3	Published in SCIE journals	Q1	30
		Q2	16
		Q3	8
		Q4	4
4	Published in EI indexed core journals (not indexed by SCIE)		3
Reward standard for academic papers in humanities and social sciences			
1	SSCI, A&HCI indexed		30
2	Authoritative university journals		8
3	Important university journals		4
4	Academic papers published in the Theory Edition of People's Daily and Guang Ming Daily or reprinted by Xinhua Digest and China Social Sciences Digest (over 2000 words) or reprinted in full by Copies of Publications of Renmin University of China (database)		4
5	Published in CSSCI regular core journals		1

Source 1) *The Notice regarding financial incentives on science and technology publication.* [EB/OL] http://kjc.cqu.edu.cn/info/1057/3920.htm; 2) *The Notice regarding financial incentives on humanity and social science publication.* [EB/OL] http://fah.cqu.edu.cn/info/1063/1183.htm. *Note* CSSCI refers to Chinese Social Sciences Citation Index. The database developed by Nanjing University since 1997 currently covers approximately 500 top Chinese academic journals of humanities and social sciences

To evaluate the quality of a paper, most universities look at the impact of journals by their inclusion in recognized indexes. Papers are usually classified in A, B, or C levels, which are weighted differently. In some universities, one A-level paper is equal to 3 C-level papers.

For scientific and engineering papers, many universities classify the *Science* and *Nature* journals as A, SCIE as B (some split SCIE into more levels according to quartiles), and EI as C. For social sciences, relatively few universities classify SSCI and the one or two most authoritative journals of a discipline as A, important journals of a discipline as B, and other CSSCI papers as C. The prominence of journals determines the size of rewards. In addition, the vast majority of universities require teachers

to be the first or sole author for the aforementioned research awards or papers. Second authorship or affiliation implies ineligibility.

In addition to direct financial rewards, the quantitative metrics in annual performance appraisals may also be linked to teachers' incomes. Some universities require teachers to produce a certain amount of research or a specific number of teaching "credits" each year. Different journals are assigned different weights to calculate research credits. In one Project 985 university, a teachers' research credit score is divided by the average score; a result below 0.5 is considered failure, whereas a result over 1 is deemed outstanding. This policy fosters a highly competitive atmosphere among teachers at various departments, because even if someone has published a substantial number of papers in objective terms, he or she may still fall short of the school average and run the risk of failing. A faculty member in the field of computer science at a Project 985 university stated in an interview that a teacher has the task to gain 2500 research credits yearly, with each 100 thousand RMB in project funding counted as 1000 credits (Interview with a professor at Computer Sciences Department 2010). In some universities, some teachers who cannot gain vertical project funding (organized and sponsored by central or local governments, wherein researchers have to compete for the funding) have to acquire horizontal project funds from an external partner (e.g., a company). Some teachers who have no horizontal projects even choose to sign research contracts with some enterprises while paying research funding out of their own pockets to meet research funding requirements (Interview with an economics teacher at a local university 2011). Under the pressure of such quantitative performance assessments, teachers are forced to increase their research output.

Bibliometric Evaluation in Promotion and Annual Appraisal

For university teachers, promotion is critical for their careers and income (Long et al. 1993). Promotion has also become an effective approach to incentivize teachers toward a large research output. At present, the majority of universities have set basic quantitative research achievement targets in academic promotion. Only when such targets are met do academics become eligible for applications. Although some universities do not have specific bibliometric evaluation in place for such appraisals, in practice, teachers with a large amount of publications tend to have a clear advantage.

In recent years, quite a few universities have reformed promotion systems and established differentiated regulations in place for research publications, particularly by discipline. In effect, a distinction is made in the appraisal standards for the natural sciences and engineering and those applied in the humanities and social sciences. However, such differentiation is often only down to the level of broad categories (e.g., natural sciences, engineering, and information sciences) and has yet to reach specific disciplines. Moreover, many universities have divided faculty into three tracks, i.e., teaching and research focused, research focused, and teaching focused. Nonetheless, mandatory research targets on research achievements invariably exist for whichever category. For example, in a Project 985 university in west China, one must meet the following requirements to apply for a full professorship position: (1) lead one national-level research project or lead research projects with 2 million RMB in accumulated funding (400,000 for humanities or social sciences); (2) publish three SCI or SSCI papers (two being in authoritative journals) or five papers in Chinese peer-reviewed journals; for social sciences, equivalent metrics are one SCI or SSCI paper or three papers in leading Chinese journals; and (3) national-, provincial-, or ministerial-level awards for outstanding achievements in research or teaching. If the candidate has not won such awards, then he or she would need another three SCI or SSCI awards (one SCI or SSCI paper or three papers in authoritative social science journals).

Most universities include external review as an important approach in the promotion process. But Reviews by external experts are often used as a mere reference. From the perspective of universities, setting bibliometric standards for promotion can save administrative costs, and avoid favoritism, thereby making appraisals fair and objective, while the disadvantages of this approach is ignored Consequently, academic council members who cast their votes don't need to face the relationship pressure and make the difficult decision on whose work is better, they can make decisions simply according to the numbers.

In some universities, the qualification to serve as a supervisor for Ph.D. students is linked to quantified research results. According to a document titled *Rules on Reviewing Ph.D. Supervisor Qualification* issued by a Project 985 university in Western China in 2005, applications to become a Ph.D. supervisor no longer require professorship but instead specific requirements regarding the number of academic publications. In the fields of natural sciences, engineering, medicine, and management, applicants

will need to have achieved "more than five papers published in authoritative international journals" as first author. Meanwhile, in humanities and social sciences, more than 10 papers in the past five years as first author in authoritative journals recognized by the graduate school are necessary. Based on our policy document analysis of Project 985 universities, many of these universities have similar policies, with some setting even higher requirements than the others.

It is found that quantitative research evaluation has found its way into every facet of university faculty evaluation. While these approaches may serve as a stimulus to increasing publications to a certain extent, they come with numerous problems. First, the evaluation methods used to determine a paper's quality are overly simplistic. The impact of a journal and whether it has been included by an authoritative database are insufficient evidence to determine the quality and impact of individual papers thoroughly. Second, over-emphasizing the quantity and efficiency of publication indicates that some basic or risky studies with long completion cycles are deprived of due attention. Third, although most universities make a distinction between the evaluation metrics for sciences and engineering and for humanities and social sciences, the inherent differences among disciplines in terms of research output and assessment are still often overlooked. Fourth, most universities consider indexes, such as SCI and SSCI, as the main standards for quality of publications. However, most of these indexes' journals are in English; therefore, some quality Chinese language journals receive less credit than they deserve. Excessive focus on publishing in English is impractical for some disciplines as well, particularly in history, philosophy, education, anthropology, and other such disciplines where considerable research output is devoted to local settings, making them less accessible to international audiences.

Having realized the flaws in such simplistic bibliometric evaluation, some universities have begun to reform their policies and systems, shifting their focus from excessive emphasis on quantity to the quality of research output. For example, Jilin University proposed in its 13th Five-Year Plan to *"explore an evaluation approach combining both quantitative and non-quantitative assessment in philosophy and social sciences, introduce a 'magnum opus' assessment system. For major achievements through committed research over an extended period, to offer retrospective and compensatory rewards and appraisal."* Fudan University, among others, also introduced a "magnum opus assessment system" some years ago and made productive efforts in non-quantitative evaluation

of research output. However, these endeavors still face formidable challenges, including the many flaws of the peer review system, which is the basis of non-quantitative evaluation. Some peer review mechanisms have faced numerous difficulties due to the lack of autonomy, intervention from administrators, and limited resources.(Zhou and Shen 2015; Jiang 2012). Wallmark and Sedig (1986) asserted that despite being overly simplistic, bibliometric evaluation has a low cost of merely 1% or less of peer reviews. In addition, the hierarchy within academic systems indicates that few elites dominate academic resources and the resource allocation process, in which politics, connections, and social capital often have a role to play, thereby compromising the fairness of peer reviews (Yan 2009).

Implications

As presented in the preceding discussion, quantitative audit, which is a systematic approach for resource allocation, includes funding, policy, and value and is now deeply inserted into every facet of universities. Government functions, such as the Ministry of Education at the macro level, universities at the meso level, and departments and researchers, have all become implementers of this system. Academics are the essential stakeholders in this system. Research audit concerns their everyday life. The number of subsidies, academic accolades, and career promotions are all determined by their quantified performance on metrics, such as quantity and quality of papers and the amount of funding. Research audit and bibliometric evaluation have initiated extensive and far-reaching reforms of scholars' academic endeavors and the academic profession as a whole, as well as a series of positive or negative consequences. On the basis of the interviews with researchers, we identify that the quantitative evaluation of research has enhanced the degree of professionalization in Chinese academia but led to the unintended consequence of "research ritualism."

Research Audit and the Professionalization of Academics

On the basis of the strong allocation function and mobilizing energy of quantitative evaluation, academic research in China, particularly in natural sciences and some engineering and social sciences, have begun accelerating their integration with the international academic community and enhancing the professionalization of academics. Professionalization herein

does not refer to the organizational establishments, that is, access thresholds of the academic profession (such as a doctoral degree) or specialized societies but rather how research paradigms, methodologies and technologies, theories and concepts, standards of academic writing, and many other processes of knowledge production have begun to be profoundly and extensively influenced by international academic norms.

The widespread use of bibliometric research metrics accompanies a wave of development of Chinese universities, with building world-class universities as a featured goal. As a state policy instrument, research evaluation has effectively guided Chinese academic research toward internationalization. After the People's Republic was established, the development of research suffered from misconceptions and detours, negatively affecting the professionalism and ethics of Chinese research academics. Quantitative research evaluation, which began toward the end of the 1980s, introduced a new idea of academic competition to the academia. Under the pressure to publish in international peer review journals, scholars (mainly in natural sciences and engineering) have to consciously improve the quality of their research to gain international recognition and publish their research results. The effect in the social sciences and humanities were complex, as argued above.

A key aspect of academic professionalization is the establishment of meritocracy and universalism in academic evaluation, which considerably changes the previously ambiguous title promotion and academic appraisal activities. Since then, the power relations of researchers and focus of their work have undergone academic shifts. Research audits have introduced forms of individualized competition, thereby transforming the human resource traditions that have been based on collectivism. Given the insignificance of peer review, non-academic standards can easily override academic considerations in the period prior to research audits. "Special factors," such as factional affiliation, seniority, position of mentor, and interpersonal relations, exist throughout a scholar's career. A scholar will need to invest energy to maintain *guanxi* (a network of relations) in the academic community in exchange for future development opportunities.

By contrast, despite its inevitable flaws, quantitative audits set clear targets and allocate academic performance a critical position. Academic performance has become the focus of academic work. Clear and universal standards have been established to determine who be promoted or rewarded, thereby eliminating social interference to the maximum extent possible. At present, the dominant logic of the system dictates that

capacity and performance will determine who will succeed, and the rules of stratification in the academia have shifted accordingly. Such a system incentivizes academic diligence and encourages scholars to focus their attention on a larger academic world rather than the complicated inter-personal relations within a department. In comparison with the previous distorted peer review systems, bibliometric evaluation is a system with more formal integrity. The president of a Project 985 university revealed that before introducing a bibliometric system, many people asked for favors before annual title appraisals. Such favors decreased after the quantitative evaluation system was set in place. Therefore, he contended that bibliometric evaluation must not be canceled lightly (President of a Project 985 university 2014).

Despite the criticism received by quantitative audits from the academia, many researchers recognize that the universalist principles of bibliometric evaluation protect them from particularism factors (See Long and Fox (1995) for details on universalism and particularism). In the Chinese academic community where academic mobility is still emerging, this universalism based on academic performance may be conducive to breaking the repression and injustice caused by a history of inbreeding and favoritism within the research community. A young humanities teacher holding a position in a rather inward-looking school, with most colleagues being graduates of that university, argued the following:

> I am a newcomer from outside. I bury my head in my research, teach my classes well, keep good terms with colleagues and that'll do for me. No need to rack my brain to play up to (leaders and colleagues). Bonuses, titles and awards require decent work. Without that, they wouldn't be able to get it, even if they had powerful people behind them. If an outsider wants to take root here, he'll need to publish papers constantly. (Assistant professor of education in one Project 985 university 2017)

Another teacher remembered her mentor's words before she graduated:

> that place (the institution she would work for) is complicated. So you bury your head in writing and publishing as many papers as possible. Put your perfectionism on hold for the moment, as long as you know where the flaws are. This will protect you... with this intense academic competition nowadays, they'll have to put some capable people to the foreground. (Assistant professor of sociology in one Project 985 university 2017)

However, some nuanced insights, as well as a long timeframe and global system support, are needed to understand how individualism based on academic performance affects the collectivism in the academic profession and how universalist principles correct particularistic principles. In fields where knowledge production is conducted on a team basis, many scholars cannot independently run a laboratory, which is relatively different from the US systems. A large research team often includes several teachers, sometimes comprising a dozen. The requisite resources, opportunities, and platforms for their careers are initially distributed within the teams, thereby affecting the academic performance of individual scholars who will need to balance academic strengths with power relations. Moreover, no simple zero–sum relationship exists between the individualistic competition and traditional collectivist culture brought in by quantitative evaluation.

Given the heightened academic competition and strengthened universalist principles, research activities now occupy a considerable amount of time and energy for teachers. Before the prevalence of bibliometric evaluation, wholehearted dedication to research came from passion and self-discipline on the part of researchers. Prior to the introduction of reward and punishment systems that oriented scholars toward research, the amount of academic pressure or time spent on research was a matter of choice for most academics, which differs from the situation in present time. An increasing number of scholars spend a considerable amount of time on academic works, especially in research universities. Numerous scholars are focusing on academic work, reflecting the macro trend of research professionalization at micro and daily levels. As one interviewee explained,

> Whenever I have time, I put it into work. Outside our world, many have the misunderstanding that university teacher is the easiest job. A few classes every week and the rest is all weekend, plus the long holidays in summer and winter. I don't know how to begin explaining this to outsiders. We never have too much off time. When there's a gap in my schedule, I'm thinking about my research. (Assistant professor of sociology in one Project 985 university 2017)

This description of this interviewee's pace at work is relatively common among research universities.

Overall, among the policies and systems in China's key university development drive, bibliometric evaluation of scholars' academic output has been the most universalist and consequential. Such policy yields an in-depth reconstruction of academic work and elevates its professionalization, which has been reflected in every facet of daily research practices.

Unintended Risk of Research Audit: Research Ritualism

While boosting professionalization, bibliometric research evaluation builds researchers a rationalized "cage." However, this approach also causes a series of unintended consequences, the most delicate yet riskiest of which is "research ritualism." Briefly, the term refers to the shift in orientation of research from knowledge to metrics. Thus, the purpose of research is not to attempt important theoretical or practical problems but to publish papers, secure projects, or obtain rewards from the system—the target number steering research. Many researchers and reviewers of the system provide little attention to the research itself and value considerably whether research will drive numbers up. Research has become a formality, void of substance and relevance, meaning rendered unimportant, and knowledge disembodied from the knowledge production scenario. Driven by the rewards and punishments of research audits, research values of "knowledge for knowledge's sake" has become "publishing for the sake of publishing." "Paper scholars" and "project scholars" are some apt terms coined to capture this shift.

Many interviewees at research universities have been aware of *"everyone going full throttle, working against the clock, not in pursuit of quality work, (but) to beat the targets."* In the critical rank advancement process, *"it (appraisal system and experts) looks at the number of papers, not at what problem a research addresses. This is a complication. This direction has very big implications"* (Interview with a professor of the Physics 2012). In some fields, such as biology,

> huge bubbles in research. 90% or papers published have no scientific significance whatsoever. They only want to publish the paper. Current evaluation system caused people to publish papers. Perhaps too few people truly want to solve a problem in science. They do exist. But there're just too few of them. (Assistant professor of biology in one local university 2017)

Quantitative auditing does not completely ignore the quality of research. However, quality, the assessment of which entails subjectivity, has been simply reduced to objective metrics—the journal in which a paper is published and its reputation. A *Cell, Nature*, and *Science* (CNS)-caliber paper is sometimes rewarded with hundreds of thousands of RMB. The system rewards the action and result of publishing in CNS or authoritative journals and not the academic value or social relevance of the research in question. This condition fosters a breed of darlings of the system; the prestige of CNS and the massive funding that comes along with it result in scholars buying expensive equipment and recruiting additional talents to publish more CNS papers. These researchers reap fortune and fame, feeling considerably at home with the research audits of the bureaucratic system. Whether they genuinely produce highly valuable research or have earned the highest honors in the international academic community have been neglected.

Research audits set meticulous metrics for bonus, promotion, and strict expiration dates for research output, which are typically the timeframe of evaluations, that is, the small annual evaluation, the big re-evaluation once every three years, the "past three years," and "past five years" in various forms, age limits in fund applications, and talent accreditations. These time management techniques build a "seize-the-moment" urgency and "now-or-never" anxiety.

Pressured by career development and financial reward, many researchers have to choose either "more" or "good" research. Compromising toward the former has become the rational option to survive in an atmosphere of audit culture. The traditional value of "it takes 10 years to properly sharpen a sword" implies immense career risks under the current system. "Your paper is out or you are out" is the destiny of all scholars on the upward curve of their academic careers. *"They only need 3 years to sharpen their sword. Maybe long before the sword is ready, they're let go", "(if) you want 10 years, you'll get no students. No students want to do it with you"* (Professor, Fudan University 2011).

In comparison with ambiguity and less interest in the old academic evaluation approach, research audits have thoroughly mobilized researchers' work commitment. "The more, the better" has become the strategy adopted by scholars facing job uncertainty with intense competition. When research evaluation is reduced to piecework, "those with the most, win" has become the first rule of survival for most faculty in the Darwinist climate of contemporary academia. "Those with the best,

win" is a less-commonly held axiom. Quantity is the prerequisite, whereas quality is the highlight. From this perspective, only by possessing both can one succeed in the fierce competition for academic resources. "Guided" by research audit policies, many researchers have developed a series of coping strategies in knowledge production. Our interviews reveal that some researchers may choose projects where results and papers can be produced in the short term. Time-consuming research works that demand undivided attention are neglected or shelved.

> He could have produced very high caliber papers but he hasn't done much of that. Working on high quality papers demands a lot of energy. If some teachers want to be practical and they want rank advancement, they will try to publish a lot of irrelevant papers. Because of publishing pressures, graduate students and Ph.D. students cannot afford to give their undivided attention to really delve into something. If they do what Chen Jingrun did, they'd have no chance of getting their degrees. To do difficult research with real value, probably time's not enough to publish a paper. (Professor of Physics 2012)

The increase in the number of authors on a paper has also been driven by quantitative research audits. Collaboration can lift a researcher's output amount, impact, or "credits." Some researchers have adopted a "dilution" tactic, splitting what can be condensed in one academic masterpiece into a few papers, *"dilute a cup of strong tea to a few cups of light tea. There are still new ideas or insights in each paper, and now your numbers are up."* (Associate Professor of Chemistry 2014)

The academic community is a highly differentiated world. Disciplinary differences exist in how social contexts influence the production of knowledge (Becher 1994). Therefore, research ritualism has various manifestations in the context of different disciplines. For example, in comparison with natural sciences in which transcending geographical differences and finding common goals across different regions are easy, the value of engineering research requires synthesis of local economy, society, and culture. Solving practical problems and promoting the integration and conversion among education, research, and industry are the key goals of many disciplines of engineering.

Current standards of quantitative research auditing mirror the standards for natural sciences. For other fields, research audits exhibit the "arrogance of ignorance" of the administrative system. Particularities of

some fields or institutions have been overlooked while setting metrics and standards. Moreover, the reliance on metrics, such as CNS and prestigious journals, impact factor, and citation frequency is squeezing out the breathing room of applied research in science and engineering subjects. A professor at an important engineering university asserted that when the entire university was fixated upon the numbers of SCI papers and citations, it ran the risk of *"losing applied disciplines."* This university had a tradition of undertaking military projects, which are often funded through the trust mechanism of commissioning due to their particularity and confidentiality. The competition mechanism of project application was rarely used, and such projects were seldom included in the statistics of "vertical projects." However, in current "research project and funding" appraisals, only vertical funding, that is, funding from national research councils, carries considerable weight. Hence, commissioned projects from state ministries, industries, and other agencies receive decreasing appreciation. As quantitative research audits delve into the fabric of academic governance, military projects at the university have diminished. Researchers are either actively or passively dragged into the selection process of the system. All their works and labor must be visible and quantifiable through the "filter" of quantitative metrics. The dominance of quantitative metrics is even more harmful to humanities and social sciences, which require extensive intellectual investment. Moreover, given that not all humanities and social science journals in China adopt a peer review system and publishing in non-peer review journals is relatively easy, some scholars tend to publish a large number of papers to gain visibility and influence.

Another symptom of research ritualism is "internationalization for the sake of internationalization." Particularly, in some social sciences and humanities disciplines, internationalization may lead to hollowness of research, detaching research contents from the Chinese context or guiding important academic questions onto biased pursuit of internationalization targets. Appraisal and competition in metrics have pushed the Chinese academia onto the world stage with their immense mobilization. The prevailing view that "international" represents superiority and is cutting-edge implies that "foreign accolades" are heavily preferred in China's research auditing. Publishing in international journals translates into considerable rewards and development opportunities. Therefore, an enormous drive to engage in internationalized research exists from the university down to academics. That is to say, "internationalization" can

become a strategy for resource acquisition rather than a genuine endeavor that stems out of the intrinsic needs of the research in question. Some schools and department prefer to recruit academics returning from overseas or place resources in fields where publishing SSCI or SCI papers is easy. Such plans may not have been conceived to build academic strongholds or enhance academic portfolios, and such new directions often struggle to synergize with traditional strengths of these schools or departments. These plans are rather designed to contribute to the internationalization metrics or discipline rankings.

Conclusion and Discussion

Bibliometric evaluation techniques have been gradually introduced to Chinese universities starting from the end of the 1980s. Among other assessments, World-Class University Rankings, as discipline appraisals, entered the domain of higher education and gained sweeping popularity, and quantitative evaluation techniques guided by publication and citation numbers gained enormous legitimacy. Analysis of university policy texts reveals that bibliometric evaluation techniques have been extensively utilized in faculty recruitment, professional promotion, reward, faculty appraisal, and other processes.

In addition to the boost provided by external assessments, the underdevelopment of the academic community and limited institutionalization of peer review culture are also key reasons for the unrestrained expansion of bibliometric evaluation. In China, although peer review has been an established practice for publishing in journals, appraising projects, reviewing academic honors, and title appraisals, rigorous implementation is often demanded or even twisted. Many journals still do not have strict peer review systems, and favoritism can be found in the publishing process. Pulling strings and favoritism are commonplace practices in the reviews of projects and academic honors. Without a strong peer review and a healthy research culture (Shi and Rao 2010), universities and scholars resort to publication numbers to improve visibility and reputation, and piecework research became the norm.

The entire academia is practically united in criticizing bibliometric research evaluation. However, in contrast to such views, this study holds that bibliometric evaluation has played some positive role in meeting the long demand of faculty professionalization due to the lag-behind of

Chinese higher education at the end of the 1970s. Ultimately, quantitative auditing of research is similar to a double-edged sword. Positive changes in China's academic work, improved professionalization, and an elevated role of research in the academic careers driven by number-driven governance approaches are observed. Meanwhile, an academic space characterized by weak academic autonomy and maldeveloped peer review also led to systemic risks most typically represented by "research ritualism." The range of choices (e.g., "10 years sharpening a sword") in academic work is gradually compressed. A large number of scholars have reminisced about the calm and composure of the good old days of lenient appraisals, reflecting the stress caused by research auditing.

The problems of bibliometric evaluation have drawn increasing attention of scholars and government officials. However, the evaluation of universities is still currently dominated by external forces that largely base their appraisals on quantified metrics. Bibliometric evaluation can well further intensify in the present new wave of Double World-Class University initiative. How Chinese scholars may respond to this bibliometric evaluation system deserves further investigation.

References

Aagaard, K., Bloch, C. W., & Schneider, J. W. (2015). Impacts of performance-based research funding systems: The case of the norwegian publication metric. *Research Evaluation, 24,* 106–17.

Anderson, M. S., Ronning, E. A., Vries, R. D., & Martinson, B. C. (2010). Extending the Mertonian norms: Scientists' subscription to norms of research. *The Journal of Higher Education, 81*(3), 366–393.

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education, 19*(2), 151–161.

Brenneis, D., Shore, C., & Wright, S. (2005). Getting the measure of academia: Universities and the politics of accountability. *Anthropology in Action, 12*(1), 1–10.

CCNU. (2016). *The notice regarding to make 2017 yearly plan.* [EB/OL]. http://hr.ccnu.edu.cn/info/1011/3137.htm.

China Institute of Scientific and Technical Information. (1995). Statistical analysis of Chinese scientific papers in 1993. *Science, 2,* 52–55. (In Chinese).

Gläser, J., & Laudel, L. (2007). The social construction of bibliometric evaluations. In R. Whitley & J. Gläser (Eds.), *The changing governance of the sciences. The advent of research evaluation systems* (pp. 101–123). Dordrecht: Springer.

Hermanowicz, J. C. (2016). The proliferation of publishing: Economic rationality and ritualized productivity in a neoliberal era. *The American Sociologist, 47*(2–3), 174–191.

Huang, Y., Pang, S. K., & Yu, S. (2018). Academic identities and university faculty responses to new managerialist reforms: Experiences from China. *Studies in Higher Education, 43*(1), 154–172.

Jiang, K. (2012). Peer review system in the construction of educational academic community. *Journal of Peking University (Philosophy and Social Sciences), 49*(02), 150–157.

Kyvik, S., & Aksnes, D. W. (2015). Explaining the increase in publication productivity among academic staff: A generational perspective. *Studies in Higher Education, 40*(8), 1438–1453.

Li, J. (2016). The global ranking regime and the reconfiguration of higher education: Comparative case studies on research assessment exercises in China, Hong Kong, and Japan. *Higher Education Policy, 29*(4), 473–493.

Li, F., Yi, Y., Guo, X., & Qi, W. (2011). Performance evaluation of research universities in Mainland China, Hong Kong and Taiwan: Based on a two-dimensional approach. *Scientometrics, 90*(2), 531–542.

Liu, W., Hu, G., Tang, L., & Wang, Y. (2015). China's global growth in social science research: Uncovering evidence from bibliometric analyses of SSCI publications (1978–2013). *Journal of Informetrics, 9*(3), 555–569.

Long, J. S., Allison, P. D., & McGinnis, R. (1993). Rank advancement in academic careers: Sex differences and the effects of productivity. *American Sociological Review, 58*(5), 703–722.

Long, J. S., & Fox, M. F. (1995). Scientific careers: Universalism and particularism. *Annual Review of Sociology, 21*(1), 45–71.

Muller, J. Z. (2018). *The tyranny of metrics*. Princeton, NJ: Princeton University Press.

Nadler, Elsa G. (1999). *The evaluation of research for promotion and tenure: An organizational perspective*. Unpublished doctoral dissertation. West Virginia University, p. 61.

National Education Commission of the People's Republic of China. (1989). *1988 Compilation of scientific and technical statistics of colleges and universities*. Wuhan: Wuhan University Press, pp. 218–219.

Paradeise, Catherine, & Jean-Claude, Thoenig. (2015). *In search of academic quality*. London: Palgrave Macmillan.

(personal interview of a president of a 985 Project University, 2014).

Qi, M. M., & Chen, T. Z. (2016). The source of the obstacles to the implementation of university development planning and the way to solve it (In Chinese). *Research in Higher Engineering Education, (04)*, 124–128.

Shi, Y., & Rao, Y. (2010). China's research culture. *Science, 329*(5996), 1128.

Shore, C., & Wright, S. (2015). Governing by numbers: Audit culture, rankings and the new world order. *Social Anthropology, 23*(1), 22–28.

Thoenig, J. C., & Paradeise, C. (2014). Organizational governance and the production of academic quality: Lessons from two top US research universities. *Minerva, 52*(4), 381–417.

Townley, B. (1997). The institutional logic of performance appraisal. *Organization Studies, 18*(2), 261–285.

Wallmark, J. T., & Sedig, K. G. (1986). Quality of research measured by citation method and by peer review—A comparison. *IEEE Transactions on Engineering Management, 4,* 218–222.

Weingart, P. (2010). The unintended consequences of quantitative measures in the management of science. In Hans Joas (Ed.), *The benefit of broad horizons. Intellectual and institutional preconditions for a global social science* (pp. 371–385). Boston: Brill.

Welch, A. (2016). Audit culture and academic production. *Higher Education Policy, 29*(4), 511–538.

Whitley, R., Gläser, J., & Engwall, L. (2010). *Reconfiguring knowledge production: Changing authority relationships on the sciences and their consequences for intellectual innovation.* Oxford: Oxford University Press.

Yichu, Shang. (1988). China's academic status in the world. *Chinese Science Foundation, 12,* 24–28.

Yan, Guang-cai. (2009). Power game and peer review system within and outside the academic community. *Education Review of Peking University, 7*(01), 124–138 (In Chinese).

Yi, L. (2011). Auditing Chinese higher education? The perspectives of returnee scholars in an elite university. *International Journal of Educational Development, 31*(5), 505–514.

Zhou, Y., & Shen, H. (2015). Peer teaching in universities: Advantages, difficulties and solutions. *Fudan Education Forum, 13*(3), 47–52.