ARTICLE

Received 29 Apr 2016 | Accepted 26 Sep 2016 | Published 25 Oct 2016

DOI: 10.1057/palcomms.2016.78

OPEN

The changing role of metrics in research institute evaluations undertaken by the Chinese Academy of Sciences (CAS)

Fang Xu^{1,2,3} and Xiaoxuan Li^{1,2,3}

ABSTRACT As an efficient S&T management tool, the Chinese Academy of Sciences (CAS) institute evaluations have experienced four major systems in which quantitative indicators have played a great role. To better understand the roles of quantitative indicators in CAS institute evaluations and institute development, this article aims to discuss the history and development of the CAS institute evaluation systems. It is concluded that institute evaluation has played a significant role not only in promoting scientific progress but also in guiding the direction and focus of the institutes' development, thereby aligning it with the CAS's strategy at different phases. Our intention is that this article offers useful insights for a wide scientific community into CAS' attitudes on the application of metrics in research institute evaluations. This article is published as part of a collection on the future of research assessment.

¹ Institute of Policy and Management, CAS, Beijing, China ² Institute of Science and Development, CAS, Beijing, China ³ University of Chinese Academy of Sciences, Beijing, China

ounded (CAS, 2016) in November 1949 in Beijing, the Chinese Academy of Sciences (CAS) is considered both the country's highest academic institution for the natural sciences and the highest science and technology (S&T) advisory body. In addition, it serves as a comprehensive R&D centre for China's natural sciences and high-tech fields. Similar to the National Academy of Sciences in the United States of America, the CAS has a board of Academies, and, similar to the International Max Planck Research School, it has its own research institutes. The CAS also has three universities. This study will focus on the CAS research institute evaluation system. The CAS now has 104 research institutes, which are distributed across 21 of China's provinces, municipalities and autonomous regions. The CAS research institutes focus mainly on fundamental studies and scientific explorations that are on the frontiers of emerging disciplines in physical, earth, environmental and life sciences and advanced technology. Currently, the CAS is staffed by a total of 63,000 employees, of which 50,000 are professionals and science/technology experts distributed among the CAS research institutes. The CAS's total revenue was approximately 450 billion vuan in 2015.

As institutional legal bodies, each CAS institute carries out independent S&T innovations and administrative work. CAS headquarters (CASHQ) is responsible for the research institutes' macro-management and makes major S&T decisions, such as appointing the institutes' directors and leadership, authorizing the institutes' strategy, evaluating the institutes, allocating resources and so on. Of all these macro-S&T management tools, institute evaluation has played a significant role not only in promoting scientific progress but also in guiding the direction and focus of the institutes' development, thereby aligning it with the CAS's strategy at different phases. On the one hand, institute evaluation, which reflects the CAS's S&T development strategy, is helpful in guiding the direction and orientation of the institutes' research activities. On the other hand, institute evaluation is an effective quality control tool to collect facts and evidence about the institutes' research performance, as well as to provide advisory opinions to CASHQ regarding S&T decision-making, such as the institutes' future development focus, resource allocations and institute directors' salaries.

The CAS initiated institute evaluation in 1993. During the past 20 years, the CAS research institute evaluations have been gradually adjusted to be compatible with the CAS's development, the institutes' characteristics and especially the CAS's strategies at different phases. Until now, CAS institute evaluations have experienced four major phases, or systems: quantitative evaluation, dual evaluation, comprehensive quality evaluation and major R&D outcome-oriented evaluation (Li and Yang, 2009). Through all those evaluation systems, the quantitative evaluation results have been both strongly and weakly adopted in S&T decisionmaking processes. To better understand the roles of quantitative indicators in CAS institute evaluations, here we discuss the development of the CAS institute evaluation systems from an historical perspective, with particular focus on the changing roles of quantitative indicators in different evaluation systems. We hope this work will inspire more ideas on the roles of quantitative indicators in evaluations in this "Metric Tide" era. This article begins with a detailed explanation of CAS institute evaluations during four phases. This is followed by a discussion focusing on the strengths and weaknesses of quantitative indicators, followed by some conclusions.

Phase I (1993-2001): quantitative evaluation system

In the 1990s, Chinese social communities introduced research evaluations, and some included the CAS as an evaluated unit.

However, because of a lack of understanding, none of the external evaluations reflected a complete and accurate picture of the CAS, and they even caused misunderstandings among the public. In this context, the CAS decided to initiate its own institute evaluations in 1993 (CAS, 1993). Given that the level of Chinese S&T was relatively low with limited resources and few research outputs, it was not the best time to conduct peer review-based research institute evaluation in 1993. Therefore, the CAS proposed a *quantitative evaluation system* to evaluate the institutes' research status and to help improve their research levels and outputs within a short time. The details of the quantitative evaluation system are shown in Table 1.

A set of quantitative indicators was fully adopted in the evaluation, and these indicators were further weighted to create a ranking score as the final evaluation result. However, the evaluation result was not related to S&T decision-making regarding resource allocation; instead, it aided each institute in understanding its research status and encouraged the institutes' research activities. The evaluation result still attracted considerable attention since it was published annually within the CAS. Researchers have had intensive arguments about the full application of quantitative indicators in the institutes' evaluation.

Phase II (1999-2004): dual evaluation system

In 1998, the State Council made a major S&T decision to build a National Innovation Systems (a Chinese national initiative), and the CAS was chosen as a pilot national institute to carry out the Knowledge Innovation Project (KIP). The overall mission of the KIP was to establish the CAS as a national natural science and high technology innovation centre with strong and continuous innovative ability by approximately 2010 (CAS KIP Evaluation Methodology Research Group, 2011). A large amount of funding was granted by the state and specifically allocated to the CAS to implement the KIP. To encourage the institutes' research activities in accordance with the KIP's mission, the CAS reformed its institute evaluation system to a *dual evaluation system*, which refers to evaluations of an institute's target completion and orientation separately (see Fig. 1) (CAS KIP Evaluation Methodology Research Group, 2012).

Specifically, an institute's scientific target completeness was evaluated by peers, and its management, task completeness and orientation were evaluated using quantitative indicators such as high quality publications, talent and major awards. Those indicators were designed to evaluate an institute's performance towards its orientation and CASHQ's macro-policy, which was to strengthen institutional innovation and cultural innovation and to encourage fundamental, strategic and forward-looking S&T contributions. Both the experts' opinions and the quantitative evaluation results were considered together to generate a final evaluation result; however, the quantitative indicators actually played a relatively greater role given that no significant differences among the experts' opinions were found (Li and Shi, 2003). The final evaluation result, which indicates institute rankings within the CAS, was directly related to S&T decision-making with respect to the institutes' overall funding allocations and the directors' annual salary.

Phase III (2005-2010): comprehensive quality evaluation system

The CAS's rebuilding work was successfully completed in the early stages of the KIP, and 2005 was the key year when the CAS entered the KIP's third phase, which was to promote the institutes' innovation capability. In this context, CASHQ reformed its institute evaluation system to a *comprehensive quality evaluation system* to promote the institutes' innovation

l able 1 Quantitative ev	aluation system	
Primary indicator	Secondary indicators	Details
Research performance evaluation	Basic research performance Applied research	Number of S&T awards, number of publications, average publications per researcher, number of total citations, and number of S&T studies (total and average). Patents applied for and authorized, S&T progress award and invention award.
Research status evaluation	perrormance Commercial performance Research work	Profits of affiliated companies, R&D investment/transfer and transformation of sci-tech achievements. Capability of undertaking national and local projects, total and average research outputs, number and amount of international cooperation projects.
	Talent Resources Infrastructure and advancec	No. of academies, percentage of researchers under 45, percentage of PhD/Master students; number of research students. Total amount of grants / average research grants per researcher, percentage of competitive grants. I research equipment, percentage of 1990s research equipment, open labs and engineering centres.

Institute's target	Scientific target completion evaluation	Reviewed by peers
completion evaluation	Management target completion evaluation	CAS policy-related indicators
	Strategic S&T tasks	40%
Institute's orientation evaluation	High-quality publications	15%
	Significant social and economic benefits	25%
	Talents development	18%
	Important rewards	2%
e.g. > Permanent staff turnover rate(5%) > Ratio of fixed funds to competitive funds > Institute-owned company's system reformation		General weights determined by CAS Party in 2001

Figure 1 | Dual evaluation system

capability and development in approximately 2004 (CAS Evaluation Group, 2007). A total of 99 research units participated in this evaluation, including 89 research institutes. Given the change in the CAS's strategy during this phase, a set of quantitative indicators was developed to observe and track the development trend of various institutes based on the previous quantitative evaluation results. In particular, an innovation capability index evaluation system was proposed as an attempt to evaluate the institutes' innovation capability in response to the CAS's strategy of *shifting the research focus to improve technological innovation ability* (Liu and Zhi, 2009). The detailed calculation of the innovation capability index is presented in online Appendix A.

The comprehensive quality evaluation system combined quantitative and qualitative information, and the quantitative indicators provided basic data support for decision processes (see Fig. 2).

Further, multi-dimensional indicators from multiple angles were adopted in this system, whereas quantitative indicators were provided as references to experts including peers from research areas, management specialists from CASHQ and directors from other CAS institutes. With the help of the innovation capability index, it was possible to ensure that the various institutes' performance, the institutes' historic and current performance, and the CAS's performance during these years were all comparable. Compared with the *dual institute evaluation system*, experts (quantitative indicators) played a greater (smaller) role in generating a final evaluation result, which was also related to S&T decision-making regarding the institutes' overall funding allocations and the directors' annual salary.

Phase IV (2011-present): major R&D outcome-oriented evaluation system

Through the successful implementation of the KIP, the CAS has accomplished its institute rebuilding effort and advanced the institutes' innovation capability. Consequently, the CAS published its future development strategy for 2020, which is known as Innovation 2020. This initiative aims to encourage the research institutes to make significant contributions in scientific and technological progress, economic and social development and national security (Lu, 2011). In this context, the twelfth 5-year plan, also known as "One-Three-Five" was proposed, which stands for *One Positioning (orientation), Three Major*

ARTICLE

Breakthroughs, and Five Key Potential Directions. One Positioning refers to how each CAS institute should specify its major research areas, unique features, core competencies, and anticipated position in international circles and should avoid homogenization with other CAS research institutes. Three Major Breakthroughs refers to major basic, strategic, and prospective S&T innovative achievements to be made in the next 5-10 years. Generally, each institute should propose no more than three breakthroughs. Five Potential Directions includes research priorities with unique features, a future competitive advantage, and potential breakthroughs; in general, each institute should establish no more than five priorities. Accordingly, CASHQ has been developing a major research outcome¹-oriented evaluation system (Bai, 2012) (see Fig. 3) since 2011, which includes expert diagnostic assessments conducted every 5 years, an overall performance evaluation held in 2015, and monitoring of key performance indicators (KPI)² to observe and track the institutes' annual research performance.

In this evaluation system, the One-Three-Five expert diagnosis assessment invites international experts to diagnose the institutes' status, advantages and disadvantages and evaluate the research quality and technical value of the main research areas to help the institutes to improve their internal management, clarify their core advantage, avoid homogenization and lay a foundation for future major innovation contributions. The performance evaluation



Figure 2 | Evaluation and decision processes in the comprehensive quality evaluation system

invites domestic experts to provide qualitative opinions on an institute's performance compared with its 5-year target, and quantitative indicators such as *funds, projects, staff information, major S&T outcome, patents, major awards and international exchanges and cooperation* are provided for the experts' reference. Notably, the performance evaluation results are directly related to S&T decision-making regarding incentive resource allocations. For example, a superior *Major Breakthrough* merits a reward of 4 million RMB.

Discussion and conclusions

From an historical perspective, the CAS research institute evaluation systems have experienced four phases, and each was proposed to be compatible with the CAS's development, the institutes' characteristics, and particularly the CAS's strategies during different phases. Within those evaluation systems, quantitative indicators have played different roles in generating final evaluation results and affecting S&T decision-making regarding resource allocations. In the quantitative evaluation system, a set of quantitative indicators was adopted fully to generate final evaluation results; however, the evaluation results were not related to S&T decision-making regarding resource allocation. In the dual evaluation system, although both the experts' opinions and the quantitative evaluation results were considered together to generate a final evaluation result, the quantitative indicators actually played a relatively large role given that no significant differences were found among the experts' opinions. Furthermore, the final evaluation result was directly related to S&T decision making in terms of the institutes' overall funding allocations and the directors' annual salary. In the comprehensive quality evaluation system, quantitative indicators were provided as references to various experts. Compared with the dual institute evaluation system, experts (quantitative indicators) played a greater (smaller) role in generating final evaluation results, which were also related to S&T decisionmaking regarding the institutes' overall funding allocations and the directors' annual salary. In the major R&D outcome-oriented evaluation system, quantitative indicators are again provided for the experts' reference. Notably, the performance evaluation results are directly related to S&T decision-making regarding incentive resource allocations. In summary, quantitative



Figure 3 | Major R&D outcome-oriented institute evaluation system

indicators played a strong role at first and then a weak role in the CAS research institutes' evaluation systems through the four phases. Further, the results of quantitative evaluation had a weak relationship at first, then a strong relationship and finally a weak relationship with S&T decision-making regarding CAS resource allocations.

From the perspective of the overall development of the CAS research institutes, the application of quantitative indicators in the CAS research institutes' evaluation systems has both strengths and weaknesses. During the time when the Chinese S&T level was relatively low, compared with nowadays technology (for example, Chinese SCI publications in 2006 is 71,000, which is 14.6 times of that in 1987. And the number of Chinese PCT patents is 5456 in 2007, which is 15 ranks ahead of that in 1997 (National Bureau of Statistics, 2008)), the application of internationally recognized quantitative indicators (such as number of SCI papers) was considered a way of introducing international peer review of Chinese research studies, which has been very helpful in improving China's S&T research quality and research level. The application of comparable international quantitative indicators has also been an effective way for Chinese research and Chinese researchers to join the international stage and synchronize with international S&T development. According to the record, the CAS's SCI papers increased from 5860 in 1998 to 12060 in 2003 (an increase of 106%). Publications in the top 20 international journals increased from 171 in 1998 to 654 in 2003 (an increase of 282%). It is reasonable to believe that quantitative evaluation has led to a great improvement in both the quantity and quality of the CAS's research outputs since the 1990s, which has helped to increase the CAS's national and international visibility and impact. In addition, the application of internationally recognized quantitative indicators has helped with the selection of potential talent within the CAS and China. Those with high potential and considerable international impact have been identified. This has ensured the development of a talent selection policy and has changed the talent structure. With the development of S&T technology, the limitations of quantitative indicators within the CAS can be observed from the following two aspects. First, by adopting one set of quantitative indicators, the CAS institutes' distinguishing features have been weakened. This could result in homogenization of the CAS institutes, which is counter to their mission and orientation. In addition, although quantitative indicators have been internationally recognized, they also cause the scientific community to focus more on quantity rather than true scientific contributions and real-life problem-solving abilities. In this context, quantitative indicators are overly restrictive, which will inevitably cause problems, especially as this perspective is unfavourable to high-quality innovative S&T achievements. Therefore, the CAS attempted to find a balance between quantitative and qualitative evaluations. An example is major innovation achievement. With the application of this indicator, those with superior performance in major innovation achievements but with unfavourable performance in the quantitative indicators would still be rewarded. The same situation can be found in the major R&D outcome-oriented evaluation system. It is the experts' responsibility to evaluate the importance and significance of an institute's major research outputs. The quantitative indicators' role in this evaluation system has been further limited to aiding the observation and tracking of the institutes' research performance and serving as references for the experts.

We also notice that, outside the CAS and the Chinese scientific community, the misuse or abuse of quantitative indicators in research evaluations has already raised international concerns. Recently, the San Francisco Declaration on Research Assessment (Raff, 2013), the Leiden Manifesto (Hicks *et al.*, 2015) and the Metric Tide (Wilsdon *et al.*, 2012) have all invoked the use of

quantitative indicators with discretion for evaluation purpose. The CAS also realized this issue a few years ago and has already reformed the orientation of its research evaluation systems to assess true scientific contribution and real-life problem solving. The major R&D outcome-oriented evaluation system is one significant example, and it adopts a much more complicated evaluation method that combines both experts' opinions and big data. Simple and mechanical number counting are neither directly adopted as evaluation results nor directly relied on for major S&T decision-making in the CAS. This reform of research evaluation not only has changed research activities within the CAS but also has been highly recognized by the wider scientific community and the central government. During the National Innovation Conference held in June 2016, President Xi (2016) gave clear instructions regarding the Reform of S&T Evaluation, which included establishing classification-based S&T evaluations that focus on S&T innovation quality, contribution and performance. The scientific, technical, economic, social and cultural value of S&T innovations should be properly evaluated. It is a challenge to use quantitative indicators wisely to evaluate high-quality and large institutes, especially when administrative intervention may sometimes result in a different balance of positive and negative effects in the use of quantitative indicators. It is a pity to avoid quantitative indicators only because they are quantitative. Further research and development on the application of quantitative indicators and how to ensure that quantitative indicators play an appropriate role in research evaluation is necessary.

Notes

- 1 Six types of major innovation contributions and innovations, including solve major scientific problems, create new research fields, achieve breakthroughs in a key technology, provide a solution, achieve remarkable social or economic benefits and provide significant and influential advice are identified as types of possible major R&D outcomes.
- 2 The key performance indicators are applied to observe and track the institutes' annual research performance.

References

- Bai CL (2012) Research institute reformation through major R&D outcomeoriented evaluation system. Bulletin of Chinese Academy of Sciences; 27 (4): 407–410, (In Chinese).
- CAS Evaluation Group. (2007) Some thoughts about CAS evaluation work. *Bulletin of Chinese Academy of Sciences*; 22 (2): 104–110, (In Chinese).
- CAS KIP Evaluation Methodology Research Group. (2011) The CAS conducting evaluation work on knowledge innovation program (during 1998-2010). Bulletin of the Chinese Academy of Sciences; 26 (2): 127–133, (In Chinese).
- CAS KIP Evaluation Methodology Research Group. (2012) Evaluation methodology for knowledge innovation program of Chinese academy of sciences. Bulletin of the Chinese Academy of Sciences; 27 (4): 485–492, (In Chinese).
- CAS. (1993) 1998 Institute Evaluation Report. Beijing.
- CAS. (2016) Introduction of CAS, http://english.cas.cn/about_us/introduction/ 201501/t20150114_135284.shtml, accessed 20 September 2016.
- Hicks D, Wouters P, Waltman L, De RS and Rafols I (2015) Bibliometrics: The leiden manifesto for research metrics. *Nature*; **520** (7548): 429–431.
- Li XX and Shi B (2003) A brief discussion about CAS institute evaluation. Bulletin of Chinese Academy of Sciences; 18 (2): 131–134, (In Chinese).
- Li XX and Yang GL (2009) Research logical model in institute evaluation. Science of Science and Management; 30 (4): 5–8, (In Chinese).
- Liu XL and Zhi TT (2009) Capacity building and knowledge innovation program in CAS. *China Soft Science*; **19** (1): 56–65, (In Chinese).
- Lu YX (2011) Thoroughly implementing innovation during the period from the present to 2020, leading and bringing along China's science and technology to realize crossover development. *Bulletin of the Chinese Academy of Sciences*; 26 (2): 121–126, (In Chinese).
- National Bureau of Statistics. (2008) Chinese achievements in economic and social development during the past 30 years' reforming and opening, http://www.gov. cn/gzdt/2008-11/13/content_1147828.htm, accessed 25 September 2016.
- Raff JW (2013) The San Francisco Declaration on Research Assessment. Journal of Experimental Biology; 2 (6): 533–534.

ARTICLE

- Wilsdon J, Allen L, Belfiore E et al (2012) The Metric Tide: Report of the independent review of the role of metrics in research assessment and management, http://dera.ioe.ac.uk/23424/2/2015_metric_tide_Redacted.pdf, accessed 15 September 2016.
- Xi JP (2016) Striving for the construction of the advanced countries in science and technology, http://news.xinhuanet.com/politics/2016-05/31/c_1118965169.htm, accessed 12 September 2016 (In Chinese).

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Acknowledgements

This work is funded by NSFC (Grant no. 71201159).

Additional information

Supplementary Information: accompanies this paper at http://www.palgrave-journals. com/palcomms

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available at http://www.palgrave-journals.com/ pal/authors/rights_and_permissions.html

How to cite this article: Xu F and Li X (2016) The changing role of metrics in research institute evaluations undertaken by the Chinese Academy of Sciences (CAS). Palgrave Communications. 2:16078 doi: 10.1057/palcomms.2016.78.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/