



Liu Zhongmin (front row, third from right) receives the prestigious First Class State Technological Invention Award from President Xi Jinping in January 2015.

EVALUATION

Moving away from metrics

The Chinese Academy of Sciences is changing the way it assesses its research. But finding the right balance in such a big organization is a tough job.

BY HUANG KUN

“I never thought it would be awarded by President Xi,” says Liu Zhongmin, a chemist at the Chinese Academy of Sciences (CAS), a country-wide network of institutions with headquarters in Beijing. “It’s a glorious moment. It is the symbol of social recognition of our research, on which my colleagues and I spent more than 30 years.”

Liu is referring to the First Class State Technological Invention Award that he received from China’s President Xi Jinping at China’s annual National Science and Technology Awards Conference on 9 January 2015. His award-winning invention is dimethyl ether/methanol to olefins (DMTO) technology, which allows the chemical industry to substitute coal for oil in many applications. That is profoundly important for China, which is poor in oil but rich in coal. Liu and his CAS

colleagues invented the process and pioneered its use in industry. The total production capacity for olefins from 7 DMTO plants in China reached 4 million tonnes per year in 2014.

Encouraging researchers such as Liu to make breakthroughs like DMTO technology is the goal of CAS’s new assessment system. “That is a good model for constructing the major outcome-oriented system for evaluation at CAS,” says Li Xiaoxuan, director of the Management Innovation and Evaluation Research Center at CAS. The centre, based in Beijing, is the academy’s official think tank for research assessment.

And CAS is big. According to its latest report, CAS has 56,000 researchers and technicians, plus another 12,000 non-technical staff. It consists of 104 research institutes across China, as well as universities, companies and other entities. In 2013, through a combination of government

funding and other sources, CAS had budget of 41.9 billion yuan (US\$6.8 billion), most of which was spent on research and development (R&D) (see page S8).

In the past, CAS has evaluated its research output with quantitative metrics and rankings. But now it is changing the way it operates to give a more nuanced appreciation of progress.

GIVING IT BACKBONE

The core of the current evaluation system at CAS is known as the One-Three-Five plan. “It’s difficult to have a panorama of everything in this big organization,” Li says. “But now many things can be connected to this backbone.”

One-Three-Five is a guide to the way that each CAS institute should operate. That is, each institute should stake out for itself one major research area where it positions its core competencies, set out to make three major breakthroughs in the next five to ten years and

map out five specific research priorities that will provide competitive advantage.

Reforms that led to the One-Three-Five plan began in 2011, when chemist Bai Chunli started his term as president of CAS. He told *Nature* at the time that his plan was to “reduce the frequency of research evaluation while improving its quality” (see go.nature.com/gajb1x). Moreover, he said, research would be judged “based on national needs and socio-economic benefits”.

The following year, Bai expanded on his ideas in an article in the *Bulletin of the Chinese Academy of Sciences* (see go.nature.com/xgnkde). In the article, Bai argued that as the growth of China’s economy slowed from the frantic pace of the past two decades, it was crucial to also shift the focus of research towards innovation. Thus, the evaluation system at CAS should encourage researchers to solve scientific problems, create new research fields and make major technological advances.

FROM QUANTITY TO QUALITY

One of the main differences between the current and past evaluation systems at CAS is the extent to which quantitative measurements are used. For much of the past decade, institutes were ranked by combining a number of indicators, each with a different weighting, in various complicated formulae. There were 24 basic indicators, including the number of publications and citations, amount of funds raised, awards gained and patents issued.

This heavily quantitative assessment approach led to what Li calls the Science Citation Index (SCI) phenomenon — an abnormal emphasis on increasing the number of papers published in journals in the SCI, even if the impact of such research on industry and society was low. Li believes that the new evaluation system will be more effective at aligning research incentives with the country’s overall needs. “The shift of emphasis to qualitative assessment will encourage major outcomes that really matter,” he says.

The quantitative system had winners and losers. “Our institute always ranked at the top,” recalls Liu proudly. He joined CAS’s Dalian Institute of Chemical Physics (DICP) in Liaoning province as a postgraduate in the 1980s. Three decades later, he is now the deputy director of DICP.

Liu believes that DICP is still one of the top five performing institutes at CAS, despite the lack of quantitative ranking. Although the basic indicators are still collected, the numbers are no longer used to calculate a rank order. Instead, each institute submits its data to CAS headquarters, which sends back a report. “That’s very useful for us to analyse variations and so adjust our work,” Liu says. These indicators now form the foundation of the One-Three-Five plan, and assessment is in the form of two sections. The first section is an expert assessment — a form of peer review



A dimethyl ether/methanol to olefins complex in the Shenhua Baotou plant in Inner Mongolia.

— to check on progress towards the strategic goal, which takes place once every five years.

The DICP’s expert assessment was conducted in 2013. The experts recognized the DMTO technology as one of the institute’s major breakthroughs, which qualified the project for additional funding from CAS. “And we have many research areas that were judged to be in the first tier in the world,” Liu adds. First tier being the equivalent of ‘world-leading’ research; there are four tiers for rating research from an international perspective.

From 2012 through 2013, CAS conducted international expert assessment of 16 research institutes, including DICP. Experts invited ranged from journal editors to chairs of international societies and associations. They assessed 90 research areas within the 16 institutes, judging 26 to be in the first tier and 53 to be in the second.

INSIDE INSTITUTES

To supplement these overarching evaluations, each CAS institute performs its own assessment of its research groups and individuals. The institutes aim to stimulate competition and ensure that the best researchers are retained.

DICP is one of the biggest institutes at CAS and employs more than 1,000 staff in around 80 research groups. DICP has been performing its own internal evaluation system for many years. In the past, DICP’s indicators included quantitative elements, but now, just as CAS is reforming to be more qualitative, DICP is also moving in that direction.

“We have cut some indicators and now are using four basic criteria: the clearness of research direction; the closeness to our One-Three-Five strategy; the progress of their work; and the potential for future development,” Liu explained. “Qualitative judgement based on these criteria comes from an academic committee — in essence, a peer review.” And these internal assessments can have high stakes.

DICP, for example, evaluates and ranks all of its research groups every two years. “The bottom 5% are eliminated,” says Liu.

Not all CAS institutes are moving as quickly or decisively as DICP. In north Beijing, near

“The shift of emphasis to qualitative assessment will encourage outcomes that really matter.”

the stadium that hosted the 2008 Olympic Games, there is a cluster of 12 CAS institutes. *Nature* talked to a scientist who works in one of the institutes, but who wishes to remain anonymous for fear of repercussions for

speaking out. He is an associate researcher in his thirties and is one of an estimated 14,000 ‘middle-class’ professionals at CAS. He says that despite these recent changes at the top level of CAS, his evaluation — and that of the many others like him — still relies heavily on the old system of indicators.

“I have received money according to the number of papers I published, with additional money for each citation of the paper,” he says. This kind of quantitative reward system is still very much in existence in many institutes, he adds, they vary only in the details. “Papers bring us monetary rewards, which are given at the end of each year and are typically equivalent to three or four months’ salary,” he says. “I have known some highly cited papers that brought the authors rewards worth more than a whole year’s salary — or even of several years’ salaries.”

These indicator numbers are also important references for career advancement within CAS. Papers, citations and funds received are considered ‘hard assets’ to show to the academic committees that evaluate individuals for promotion. Chinese society values relationships — knowing the right people is important and the academic circle is no exception. Therefore, relying on hard assets is one way to make sure that evaluations are not unduly influenced by



China's President Xi Jinping (third from left) visits the Chinese Academy of Sciences in Beijing in 2013.

scientists' friendship circles. So there is certain rationality in utilizing these numbers, leaving the balance between quantitative and qualitative methods to be a delicate issue.

"Inviting foreign experts to participate in the assessment section of the One-Three-Five plan is a good attempt to solve this issue," says Mao Shude, an astrophysicist at the National Astronomical Observatories of the Chinese Academy of Sciences (NAOC), based in Beijing, who spent 22 years studying and working in the United States and Europe. Mao explains that a group of high-profile international experts will provide a more objective evaluation as they have fewer conflicts of interest or established relationships than domestic ones.

ROOM FOR IMPROVEMENT

Mao, who recently took a joint position at Tsinghua University in Beijing, says that the reform of the evaluation system at CAS is "on the right course". But compared with similar processes at international institutions, he says, there is still room for improvement.

"At foreign universities, the contribution of a person consists of three parts," only one of which is their research, says Mao. The second part is teaching, which has been largely absent at CAS but is becoming more important since the University of the Chinese Academy of Sciences started recruiting undergraduates in 2014. The third part, Mao says, is scientific cooperation and the research environment. "Groups within CAS institutes tend to operate separately, which is not conducive for an interactive and cohesive atmosphere."

Mao returned to China in 2010 as part of the national '1,000 Talent Plan', which recruits high-level experts — both Chinese nationals and foreign professors — from overseas and offers them well-funded academic positions to carry out their research in China. But, he

says, opportunities at Chinese institutions should be based on academic merit, not just experience gained overseas. "Although I am a beneficiary of this system, it's not fair for some domestic researchers who did similar or even higher-quality work but were not recognized as such." This situation is changing, however, as the Chinese government has just started a similar talent programme (the '10,000 Talent Plan') for domestic researchers, to convince them to stay in China rather than relocating overseas, notes Mao.

Another area for improvement is the way in which the evaluation considers the specific needs of female researchers, says Mao. He notes that the percentage of females in permanent positions in his division, from assistant researcher to the highest levels, is just below 15%. "I have two daughters and I am very sensitive about gender equality," he says. "It should be taken into account that women often bear additional family responsibilities."

In the Gravitational Lensing and Galaxy group at NAOC, where there are more than a dozen researchers, Wang Yuting is the only woman. Although she feels she is treated as equal to her male colleagues, she is aware that most female postdoctorates like her are at an age at which they feel they have to choose between research and family life. "Some female scientists did give up research because of family and kids."

According to the United Nations Educational, Scientific and Cultural Organization, 30% of the world's science researchers are women. Although specific data are not available, the situation is broadly similar in

"Groups within CAS institutes tend to operate separately, which is not conducive for an interactive atmosphere."

China and is more acute in the more senior positions. In 2011, the government set a target of increasing the percentage of women in all high-level professional positions to 35%, including a special section devoted to improving the number of women as career scientists. For this to be achievable, the evaluation system would need to be more flexible, as it is in many Western countries, to allow for a career break and support when women return after childbirth, says Mao.

MORE LEADS

With regards to determining the trajectory of CAS, the most important voice is that of China's President Xi. The president visited CAS in July 2013 and in a speech suggested that CAS should aim for "four leads": the development of technology; nurturing the best minds; contributing to science and technology policy; and becoming a first-class international research organization.

In 2014, based on Xi's instructions, CAS started a project called Lead Action. It proposed that its institutes be divided into four categories, each with different evaluation systems based on their missions. These are: institutes that aim to meet the needs of national strategy and industrial development should be evaluated by the people and organizations who use the technologies; institutes that focus on academic achievements should be evaluated by an international academic peer group; institutes with big scientific facilities should be evaluated by users and colleagues; and other institutes with subject specialities should be evaluated by experts in the same field.

Lead Action is gathering momentum. In February 2015, CAS released its latest guidelines for the Academy (the last revision was in 2002). Alongside the four leads are three orientations to ensure that CAS is matching its output with global science and technology efforts, the major needs of the country; and the "main battlefields of the national economy".

The outlook of the evaluation system at CAS is evolving. There are still issues to address in the way that different groups are treated, in ensuring that the various institutes match the overall strategy, and in finding the balance of qualitative and quantitative measurements. But the direction is clear. As Liu explains, flaws in the former evaluation system gave unclear signals to scientists as to what was the most valuable outcome. "There were multiple batons that pointed in different directions," he says. Now, he believes that the modified assessment system has brightly illuminated the goal: "Chinese scientists should do things that are useful to China first of all, then they could consider the rest of the world." ■

Huang Kun is a science correspondent based in Beijing.